

# Cepheids in Open Clusters: An 8-D All-sky Census

Richard I. Anderson<sup>1\*</sup>, Laurent Eyer<sup>1</sup>, and Nami Mowlavi<sup>1</sup>

<sup>1</sup> *Observatoire de Genève, Université de Genève, 51 Ch. des Maillettes, CH-1290 Versoix, Switzerland*

21 December 2012

## ABSTRACT

Cepheids in open clusters (cluster Cepheids: CCs) are of great importance as zero-point calibrators of the Galactic Cepheid period-luminosity relationship (PLR).

We perform an 8-dimensional all-sky census that aims to identify new *bona-fide* CCs and provide a ranking of membership confidence for known CC candidates through membership probabilities. The probabilities are computed for combinations of known Galactic open clusters and classical Cepheid candidates, based on spatial, kinematic, and population-specific membership constraints. Data employed in this analysis are taken largely from published literature and supplemented by a year-round observing program on both hemispheres dedicated to determining systemic radial velocities of Cepheids.

In total, we find 13 *bona-fide* CCs, 3 of which are identified for the first time, including an overtone-Cepheid member in NGC 129. Inconclusive cases are discussed in detail, some of which have been previously mentioned in the literature. Our results are inconsistent with membership for 7 candidates that have been studied previously. We employ our *bona-fide* CC sample to revisit the Galactic PLR and obtain results consistent with most other calibrations, being limited by cluster uncertainties.

In the near future, Gaia will enable our study to be carried out in much greater detail and accuracy, thanks to data homogeneity and greater levels of completeness.

**Key words:** methods: data analysis, catalogs, astronomical data bases: miscellaneous, stars: variables: Cepheids, open clusters and associations: general, distance scale

## 1 INTRODUCTION

The search for Cepheids in Galactic open clusters (CCs) has been a topic of interest in astronomy for the past 60 years, owing largely to their importance as calibrators of the Cepheid period-luminosity relation (PLR), discovered a century ago among 25 periodic variable stars in the SMC by Leavitt & Pickering (1912).

The proportionality between the logarithm of Cepheid pulsation periods and their absolute magnitudes, i.e., their luminosities, gives access to distance determinations and has established period-luminosity relationships as cornerstones of the astronomical distance scale (e.g. Freedman et al. 2001; Sandage et al. 2006). For reviews on Cepheids as distance indicators, cf. Feast (1999); Sandage & Tammann (2006), for instance.

The existence of the Cepheid PLR is most obvious among Cepheids in the Magellanic Clouds (e.g. Udalski et al. 1999; Soszynski et al. 2008, 2010), due to common distances (small dispersion), large abundance (thousands), and relative proximity (detectability). However,

knowledge of the zero-point(s) of such relations is also required; in this case, the distances to the Magellanic Clouds. For such zero-point calibrations, PLR-independent distance estimates are required, e.g. from trigonometric parallaxes (Feast & Catchpole 1997; Benedict et al. 2007), Baade-Wesselink-type methods (Gieren et al. 1997; Storm et al. 2011), or objects located at comparable distance, e.g. water masers (Macri et al. 2006) or open clusters (Turner et al. 2010).

For open clusters, distances can be determined via zero-age Main Sequence or isochrone fitting. If membership can be assumed at high confidence, the cluster provides the independent estimation of the Cepheid’s distance. Confidence in cluster membership is thus critical for such calibrations.

Since the first discovery of CCs by Irwin (1955, identified S Nor in NGC 6087 and U Sgr in M 25) and Feast (1957, established membership via radial velocities), many researchers have contributed to this field, e.g. van den Bergh (1957); Efremov (1964); Tsarevsky et al. (1966); Turner (1986); Turner et al. (1993); Baumgardt et al. (2000); Hoyle et al. (2003); An et al. (2007); Majaess et al. (2008); Turner (2010). Nevertheless, relatively few *bona-fide* CCs (< 30) have thus far been discovered.

\* E-mail: richard.anderson@unige.ch

We therefore carry out an all-sky census of classical Cepheids in Galactic Open Clusters that aims to increase the number of *bona-fide* CCs and allows us to rank confidence in membership according to membership probabilities. Our approach is 8 dimensional in the sense that 3 spatial, 3 kinematic, and two population parameters (iron abundance and age) are used as membership constraints. Both data inhomogeneity and incompleteness are critical limitations to this work, and are acknowledged in the relevant sections. We describe our analysis in Sec. 2.

For the first time, we systematically search for cluster members among Cepheid candidates from surveys such as ASAS, NSVS, ROTSE, and also from the suspected variables in the General Catalog of Variable Stars. Most data employed to do so are taken from published catalogs or other literature. However, we also perform radial velocity observations of Cepheids on both hemispheres and determine systemic velocities,  $v_\gamma$ . To improve sensitivity to binarity, literature RVs are added to the new observations. The data compilation is described in Sec. 3.

The *bona-fide* CCs identified by our analysis are presented in Sec. 4.1 and certain cluster-Cepheid combinations (Combos) are discussed in some more detail. A full table containing the cluster Cepheids candidates identified in this work is provided in digital form through the CDS<sup>1</sup>. Combos that deserve observational follow-up are identified in the text. Particular attention is given to Combos previously discussed in the literature. Discussions of additional Combos can be found in the online appendix. In Sec. 4.2, we then employ our *bona-fide* CC sample in a calibration of the Galactic Cepheid PLR. The method and results are discussed in Sec. 5, which is followed by the conclusion in Sec. 6.

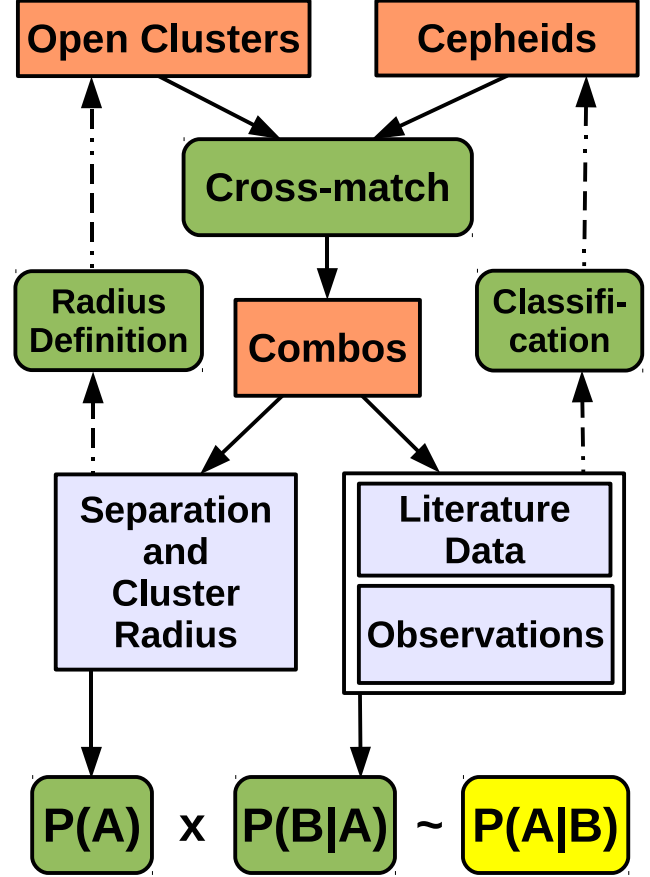
## 2 MEMBERSHIP ANALYSIS

Our all-sky census is structured as shown in Fig. 1. First, lists of known open clusters and known Cepheid candidates are compiled, see Sec. 3 for details. Second, the two lists are cross-matched positionally in a many-to-many relationship so that we investigate a given Cepheid’s membership in multiple different open clusters, and a given open cluster can potentially host multiple Cepheids. The correct classification of cross-matched Cepheid candidates is verified by considering light curves, and spectra. Misclassified objects are removed from the Cepheids sample. Third, membership probabilities are calculated based on all available membership constraints. These last two points are described in the present section.

Membership probabilities are calculated following Bayes’ theorem that can be formulated as (Jaynes 2003, § 4):

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \propto P(B|A) \times P(A). \quad (1)$$

The posterior probability  $P(A|B)$  (‘membership probability’) is proportional to the product of likelihood,  $P(B|A)$ , and prior,  $P(A)$ .  $P(B|A)$  represents the conditional probability of observing the data under the hypothesis of membership, and  $P(A)$  quantifies the degree of initial belief in membership. The normalization term  $P(B)$ , of which we possess



**Figure 1.** Schematic view of membership analysis. Rectangular boxes represent data sets used, green rounded boxes indicate actions. Cepheids are cross-matched with open clusters to form Combos. Only open clusters with consistently defined core and limiting radii are considered. Data from the literature and new observations are combined for each Combo. Cepheid classification is verified based on the data compiled (light curves, spectra). Priors,  $P(A)$ , and likelihoods,  $P(B|A)$ , are calculated separately and joined as membership probabilities,  $P(A|B)$ .

no knowledge, is the probability to observe the data. We define  $P(A)$  in Eqs. 3 & 4 and  $P(B|A)$  in Eq. 7 below.

### 2.1 Prior Estimation and Positional Cross-match

#### 2.1.1 Positional Cross-match

On-sky proximity is a necessary, but insufficient criterion for membership. Intuitively, if no other information is available, one might tentatively assume membership for a Cepheid that falls within the core radius of a potential host cluster.

Therefore, our census starts with a positional cross-match that aims to identify all combinations of cluster-Cepheid pairs that lie sufficiently close on the sky to warrant a membership probability calculation (Combos). The cross-match itself is straightforward: if the separation between a cluster’s center coordinates and the Cepheid’s coordinates is less or equal to 5 limiting cluster radii<sup>2</sup>, we include the Combo in our analysis. Using this proximity criterion, we

<sup>1</sup> cds.u-strasbg.fr

<sup>2</sup> This cut-off radius was adopted to include possible members of

cross-match 583 different open clusters (from 1658 initially compiled) with 1028 Cepheids (from 1844 initially compiled) and obtain 2570 Combos that we investigate for membership.

The initial cross-match is purely positional, and the majority of Combos studied are non-members. Our analysis intends to weed out this majority and to indicate to us the good candidates through a high membership probability.

### 2.1.2 The Prior

We define the prior,  $P(A)$ , using the on-sky separation<sup>3</sup> between cluster center and Cepheid, weighted by the cluster radius, i.e. its apparent size on the sky.

The radius of an open cluster is typically determined by fitting an exponential radial density profile to a stellar over-density on the sky, an approach originally developed for globular clusters by King (1962). The method relies on the assumption that two separate distributions are seen: a constant field distribution and one that is attributed to the cluster.

Various ways to define cluster radii can be found in the literature. Among these are the ‘core radius’ (most stars belong to cluster),  $r_c$ , and the ‘limiting radius’ for the cluster halo (strong field star contamination),  $r_{\text{lim}}$ , see Kharchenko et al. (2005b, from hereon: K05) and Bukowiecki et al. (2011, from hereon: B11).

Intuitively, the probability of membership is related to separation and cluster radius, cf. Sánchez et al. (2010). Let us therefore define the quantity  $x$  as:

$$x = \frac{r - r_c}{2r_{\text{lim}} - r_c}, \quad (2)$$

where  $r$  denotes separation.  $x$  is negative, if the Cepheid lies within the cluster’s (projected) core and becomes unity at a separation equal to twice the limiting radius. We define our prior,  $P(A)$ , so that (no other constraints considered) membership is assumed when the Cepheid lies within the cluster’s core, i.e.  $x < 0$ . Outside  $r_c$ , inspired by radial density profiles of star clusters, we let the prior fall off exponentially and define it to reach 0.1percent =  $10^{-3}$  at  $x = 1$ . Hence,

$$P(A)(x < 0) \equiv 1 \quad (3)$$

$$P(A)(x \geq 0) \equiv 10^{-3x}. \quad (4)$$

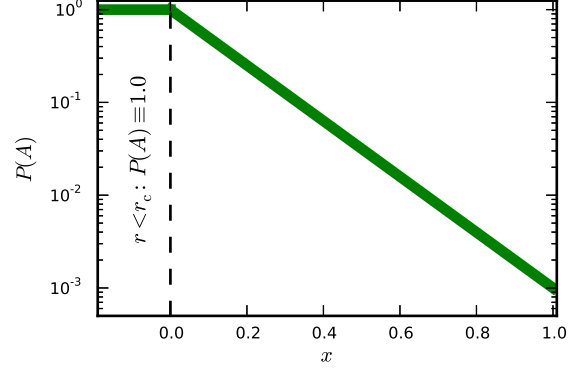
Figure 2 serves to illustrate this definition. The prior thus carries the 2-dimensional information of separation and cluster radius, and thereby takes into account how concentrated a cluster is on the sky assuming circularly distributed member stars.

## 2.2 The Likelihood $P(B|A)$

The likelihood,  $P(B|A)$ , is computed as a hypothesis test. It estimates the probability that the observed data is consistent with the null hypothesis of (true) membership. This

cluster halos in the analysis, inspired by the well-known case of SZ Tau in NGC 1647.

<sup>3</sup> We avoid the term ‘distance’ when referring to the on-sky separation (in arcmin) in order to prevent confusion with radial distance (in pc)



**Figure 2.** Illustration of the adopted prior,  $P(A)$ , as a function of separation normalized to a cluster’s radii, expressed by the quantity  $x$ , cf. Eq. 2. If  $x \leq 0$  (Cepheid within core radius,  $r_c$ ):  $P(A) \equiv 1$ . Outside the core,  $P(A)$  decreases exponentially, inspired by radial density profiles of stars clusters. We adopt  $P(A) = 0.001$  at  $r = 2r_{\text{lim}}$ .

approach was inspired by the Hipparcos astrometry-based studies by Robichon et al. (1999) and Baumgardt et al. (2000). We extend it here to take into account up to 6 dimensions using parallax,  $\pi$ , radial velocity (RV), proper motion,  $\mu_\alpha^*$  and  $\mu_\delta$ , iron abundance,  $[\text{Fe}/\text{H}]$ , and age (open clusters assumed to be co-eval), weighting all constraints equally.

Assuming that a given Cepheid was not used to determine a cluster’s (mean) parameters, we can calculate the quantity

$$c = x^T \Sigma^{-1} x, \quad (5)$$

where  $x$  denotes the vector containing as elements the differences between the (mean) cluster and Cepheid quantities:

$$x = (\pi_{\text{Cl}} - \pi_{\text{Cep}}, \langle v_{r,\text{Cl}} \rangle - v_{r,\text{Cep}}, \dots). \quad (6)$$

Let  $\mathbf{C}_{\text{Cl}}$  be the covariance matrix of the cluster and  $\mathbf{C}_{\text{Cep}}$  that of the Cepheid. Let  $\Sigma$  then denote the sum of the two and  $\Sigma^{-1}$  its inverse. Since the data employed in this calculation comes from many different sources, no knowledge of correlations between the different parameters is available. We thus make the assumption of independent measurements, which results in diagonal covariance matrices containing only parameter variances. Possible correlations between Cepheid and Cluster parameters are thus assumed to be negligible. We consider this justified, since we possess no knowledge of the extent of such correlations and assume that Cepheids were not used in the determination of cluster mean values. This formulation furthermore implicitly assumes normally (Gaussian) distributed errors.

Under these assumptions  $c$  is  $\chi^2$  distributed, i.e.  $c \sim \chi^2_{N_{\text{dof}}}$ , where  $N_{\text{dof}}$  is the number of degrees of freedom equal to the length of vector  $x$ , ranging from 1 to 6.  $c$  thus depends on the number of membership constraints considered (the on-sky position is used in the prior). In cases where no membership constraints are available, i.e.  $N_{\text{dof}} = 0$ , we set  $P(B|A) \equiv 1$ .

$P(B|A)$  is obtained by calculating unity minus the p-value of  $c$ ,  $p(c)$ :

$$P(B|A) = 1 - p(c). \quad (7)$$

Since the  $\chi^2$  distribution (and therefore the p-value com-

puted) is very sensitive to  $N_{\text{dof}}$  for small  $N_{\text{dof}}$ ,  $P(B|A)$  naturally contains information on the number of membership constraints employed.

Of course, we cannot prove the null hypothesis, only exclude it. However, by including the greatest number of the most stringent membership constraints possible, this method very effectively filters out non-members. The remaining candidates can therefore be considered *bona-fide* members, provided the constraints taken into account are sufficiently strong.

The filtering effectiveness of the likelihood strongly depends on the uncertainties adopted for the constraining quantities: the larger the error, the weaker the constraint. Conversely, the smaller the error, the more important become systematic differences between quantities measured or inferred through different techniques. Obtaining reasonable estimates of the external uncertainties is of paramount importance to the success of this work, since the data considered is inhomogeneous and listed uncertainties typically provide formal errors or estimates of precision.

For certain quantities, we therefore adopt increased error budgets that we motivate and detail in the following sections. Care is taken to avoid too large or too small error budgets, and to ensure that likelihood remains an effective membership criterion.

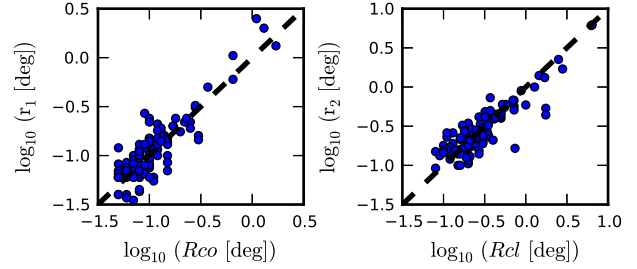
### 3 DATA USED TO COMPUTE LIKELIHOODS

In this section, we describe how we compile the data used for our analysis. The constraints employed are: on-sky separation, parallax, proper motion, radial velocity, and the population parameters iron abundance (as a proxy for metallicity) and age. Most data considered originates from published literature and catalogs. However, we also include radial velocity (RV) data from an extensive, year-round observation program carried out on both hemispheres. Some details on this program are provided in Sec. 3.2.3.1. A full description, however, is out of scope for this work and will be published separately.

Very often, data on a given membership constraint can be found in different references. In such cases, a choice of which reference to prefer over the other ones has to be made. In each of the following subsections, the references mentioned first are the ones preferentially adopted. This section is divided into two parts: Sec. 3.1 dedicated to open clusters, and Sec. 3.2 to Cepheids.

#### 3.1 Open Cluster Data

The definition of  $P(A)$  in Eq. 4 sets an important constraint on the literature on open clusters to be considered, since both a core and limiting radius have to be defined and both quantities have to be comparable in nature between different catalogs. We found three extensive studies that include large numbers of open clusters and satisfy this requirement: K05, B11, and Kharchenko et al. (2012) (from hereon: K12). The study by Froebrich et al. (2007) was not included here, since the false positive rate may exceed 50 per cent (see also Camargo et al. 2010), and since we noticed a suspicious correlation between  $r_c$  and  $r_{\text{lim}}$ . Note, however,



**Figure 4.** *Left panel:* comparison between  $r_1$  in K12 and the core radius,  $R_{co}$ , in K05; *right panel:* same for  $r_2$  in K12 and the limiting radius,  $R_{cl}$ , in K05. The radii are comparable.

that some clusters listed in K12 were originally identified by Froebrich et al. (2007).

Since many clusters in K12 were also studied by K05, we compare the three radii defined in K12 with the core and limiting radii in K05 and notice that the limiting radius in K05,  $R_{cl}$ , corresponds well to  $r_2$  in K12 (though  $r_2$  tends to be smaller), while  $r_1$  in K12 is rather similar to the core radius in K05,  $R_{co}$ . Nevertheless, a fair amount of scatter exists between both studies, see Fig. 4. We consider K12 an update (and extension) of K05 and therefore prefer the newer cluster parameters over the older ones.

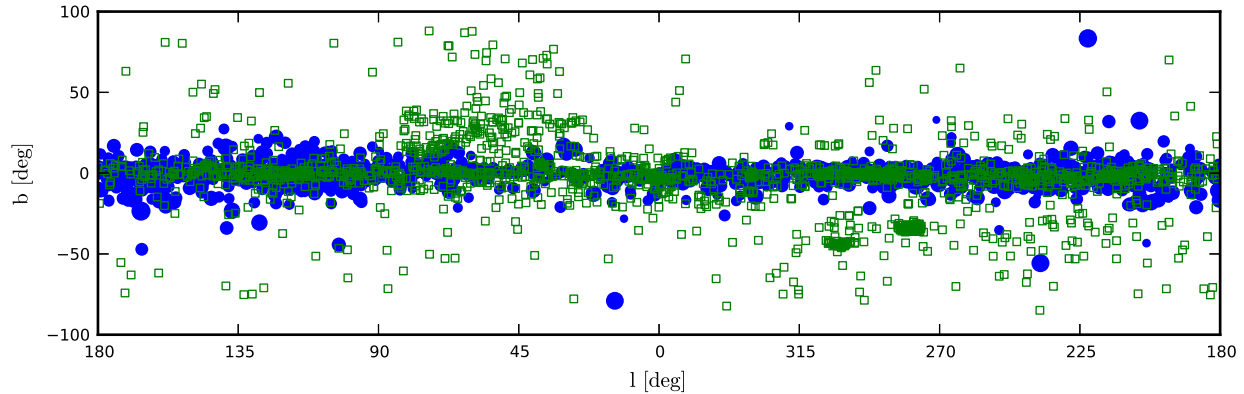
We previously compared radii given in K05 and B11 for the clusters common to both works in Anderson et al. (2012). Rather large scatter is present (more than a factor of 2 for an appreciable fraction) and illustrates that cluster radii are subject to significant uncertainty. However, the radii from both studies follow the same trend and we therefore consider them comparable for our purpose, although K05 is based on optical and B11 on near-infrared (NIR) 2MASS (Cutri et al. 2003) photometry.

Given the sometimes rather large difference between cluster radii mentioned in the literature, we adopt a ‘permissive’ scheme that gives preference to the study giving the largest limiting radius for the cluster in question in order to take into account the most interesting cases (in number and qualitatively speaking) and thereby bias ourselves towards higher  $P(A)$ . We therefore strongly rely on the remaining membership constraints that define the likelihood of filtering out chance alignments.

We then add to our compilation data from the Dias et al. (2002) catalog<sup>4</sup> (from hereon: D02), the extended Hipparcos catalog (from now on: XHIP) by Anderson & Francis (2012) and the open cluster parameters catalog by Glushkova et al. (2010). Following D02, we remove the clusters Haffner 3, 5, 25 and Hogg 3 from our sample, since they are probably not clusters (Piatti et al. 2011). The final list contains 1658 entries, 1296 of which are open clusters that are also listed in D02, 7 of which are OB associations from K05, and 355 of which were newly added by including K12. 583 open clusters are cross-matched with Cepheids and evaluated as potential Cepheid hosts.

Unfortunately, half of the 24 Cepheid-host clusters mentioned in Turner (2010) are not present in the catalogs used

<sup>4</sup> Version 3.2 as of 26 Jan 2012



**Figure 3.** Distribution of open clusters and Cepheids compiled, shown in Galactic coordinates. Blue solid circles represent open clusters with markers logarithmically scaled for apparent size. Green open squares represent Cepheids.

to compile our list of clusters<sup>5</sup>. We therefore cannot investigate these particular cases here.

Figure 3 shows the distribution of clusters (blue solid circles, scale with limiting radius) and Cepheids (green open squares) in Galactic coordinates. Clusters closely trace the disk, and no obvious gaps are present in our all-sky census.

As mentioned in Sec. 2.1, we adopt a ‘permissive’ scheme regarding the open cluster radii. We consistently apply this logic and always adopt the center coordinates and radii of the reference giving the largest limiting radius. For consistency, the parameters distance, reddening, age, etc. are preferentially taken from the same source. Distances in [pc] (and their uncertainties) are converted to parallaxes in [mas] through Eqs. 12 and 13, see Sec. 3.2.1.

In the following subsections, we outline how we construct the open cluster data set for proper motion, radial velocity, iron abundance, and age.

### 3.1.1 Mean Proper Motion

Mean cluster proper motions,  $\bar{\mu}_{\alpha, Cl}^*$  and  $\bar{\mu}_{\delta, Cl}$ , are taken from K05, K12, and D02.

The uncertainties on mean proper motion listed in these references are typically calculated either as intrinsic dispersions (for clusters closer than approx. 400pc, K12), or as standard mean errors, i.e. the error decreases as  $\sqrt{N_* - 1}$ , where  $N_*$  is the number of stars considered members, cf. D02<sup>6</sup>. The quoted uncertainties on the cluster mean are thus much smaller than the uncertainty on an individual cluster star’s measurement. For example, in K12, the typical mean proper motion error is  $0.4 \text{ mas yr}^{-1}$ .

For the majority of Cepheids, however, the uncertainties on proper motion are much larger, and many have been obtained from different data sets, using different techniques. Therefore, to ensure comparability of inhomogeneous data and to reduce our sensitivity to offsets in zero-points due to data-related specificities such as reduction techniques,

we adopt a more generous error budget for  $\bar{\mu}_{\alpha, Cl}^*$  and  $\bar{\mu}_{\delta, Cl}$  that resembles the uncertainty of an individual cluster star’s proper motion. This is done by multiplying the uncertainty listed by the factor  $\sqrt{N_* - 1}$  and thus slightly reduces the weight of proper motion as a membership constraint. Empirically, we are confident that this is justified, since proper motions of Cepheids typically barely exceed their uncertainties, and care should be taken not to over-interpret their accuracy.

### 3.1.2 Mean Radial Velocity

Average radial velocities for open clusters were taken from K05, K12, the extended Hipparcos compilation (XHIP), and D02. We adopted the XHIP values where available, and otherwise adopt the mean cluster velocity based on the largest number of stars<sup>7</sup>.

Qualitative differences exist in the uncertainties listed for mean cluster RVs. For some well-studied clusters, the uncertainty given is an estimate of the intrinsic RV dispersion. The majority of cluster RVs, however, were estimated from only a few stars (about half on two stars or less, cf. Fig. 5) and are therefore subject to systematic uncertainties due to implicit membership assumptions, for instance. In addition, unseen binary companions and instrumental zero-point offsets can introduce systematic uncertainties at the level of a few  $\text{km s}^{-1}$ .

We therefore adopt  $2 \text{ km s}^{-1}$  as a minimum uncertainty of the mean cluster velocity. If no uncertainty estimate is given, we adopt  $\sigma(RV_{Cl}) = 10 \text{ km s}^{-1} / \sqrt{N_{RV}}$  as a typical uncertainty on the mean cluster RV, where  $N_{RV}$  is the number of stars used to determine the mean cluster RV.

### 3.1.3 Iron Abundance

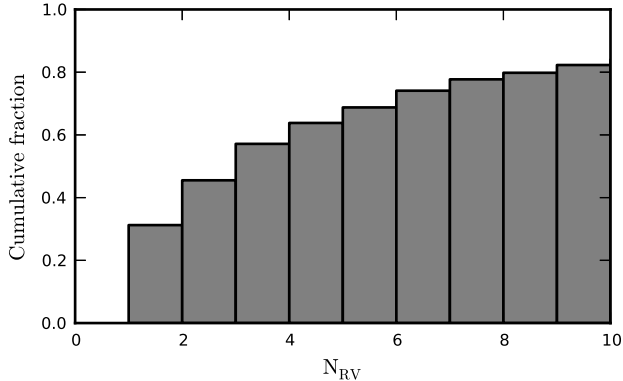
We adopt iron abundances compiled in D02, including the uncertainties given. The mean uncertainty among the clusters compiled is 0.08 dex.

<sup>5</sup> The missing clusters are: ADS 1477, ADS 5742, Anon Vul OB, Cas R2, Dolidze 45, Lyngå 6, NGC 6649, Ruprecht 175, Trumpler 35, Turner 2, Vel OB5, Vul OB1

<sup>6</sup> See under ‘version 2.3 (25/abr/2005)’ in file: <http://www.astro.iag.usp.br/~wilton/whatsnew.txt>

<sup>7</sup> With the exception that the values in K12 replaced the values published in K05





**Figure 5.** Cumulative fraction of clusters for which a given number of stars,  $N_{RV}$ , was used to determine the average RV. About half of average cluster RVs are based on a couple of stars.

### 3.1.4 Cluster Age

Ages were available for most clusters, since they are often determined simultaneously with the distance via isochrone-fitting. Such was the case in B11, for instance, or for many clusters in K05. Although a model-dependent parameter, age does provide a valid constraint for membership, reflecting evolutionary considerations that are empirically validated. Quantifying an uncertainty for age as a parameter, however, is rather difficult.

Younger clusters exhibit a Main Sequence turn-off at higher stellar masses than older clusters. As a consequence of the Initial Mass Function, a younger cluster's turn-off point tends to be less populated than that of an older cluster. It therefore follows that age estimates tend to become more accurate with age, since the cluster's turn-off point tends to be defined more clearly against the field and therefore better constrains an isochrone fit.

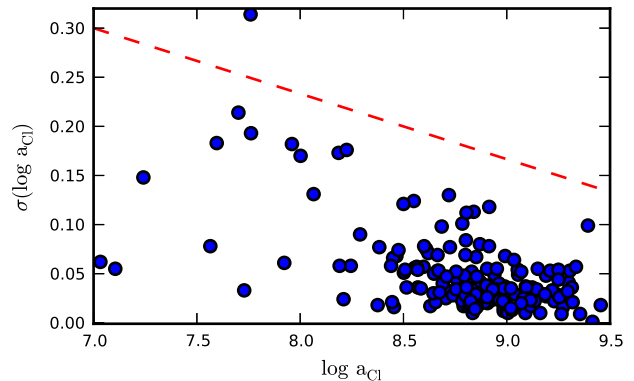
Figure 6 corroborates the above reasoning by showing cluster ages against their uncertainties as given in K12. We thus estimate an upper limit on the uncertainty of cluster age as the red dashed line in Fig. 6, which is:

$$\sigma(\log a_{CI}) \leq 0.3 - 0.067 (\log a_{CI} - 7.0) . \quad (8)$$

## 3.2 Cepheid Data

Cepheid candidates were compiled from the January 2012 version of the General Catalog of Variable Stars (Samus et al. 2012, from hereon: GCVS) and the May 2012 version of the AAVSO Variable Star Index (from hereon: VSX)<sup>8</sup>. From GCVS and VSX, we import the variability types CEP, CEP(B), DCEP, DCEPS; from VSX, we include the ASAS Pojmanski (1997, 2002); Pojmanski et al. (2005) Cepheid candidates classified as DCEP-FU or DCEP-FO. This list also contains Cepheid candidates found by ROTSE (Akerlof et al. 2000) or NSVS (Woźniak et al. 2004), as well as the ones in the suspected variables catalog (Kukarkin & Kholopov 1982).

This starting point contains an unknown, but probably



**Figure 6.** Age uncertainty,  $\sigma(\log a_{CI})$ , as function of cluster age,  $\log a_{CI}$ , given by K12. Older clusters have more precisely estimated ages. We adopt as error budget for clusters without stated age uncertainties an upper limit to this proportionality indicated by the red dashed line, cf. Eq. 8.

high, fraction of non-Cepheids. Contamination due to type-II Cepheids, or Cepheids belonging to the Magellanic Clouds are removed from the sample during the cross-match with the clusters (trace the disk). To further reduce contamination, we visually inspect all ASAS-3 V-band light curves of Cepheid candidates with ASAS identifiers.

Radial pulsation and color variations during the pulsation are defining characteristics of Cepheids. We thus use the spectra obtained for radial velocity observations described in Sec. 3.2.3.1 to verify classification. A total of 146 ASAS Cepheid candidates and 18 others are thus rejected from the Cepheid sample, resulting in a final list of 1844 Cepheid candidates, 1028 of which are cross-matched with open clusters.

The cleaned sample of Cepheids cross-matched with clusters was appended with literature data from many sources, and references are given in the text. Among the most relevant references are:

- The Fernie et al. (1995) DDO Cepheid database<sup>9</sup>
- The Klagyivik & Szabados (2009) Cepheid database (KS09)
- The ASAS Catalog of Variable Stars (Pojmanski et al. 2005, ACVS) and associated photometry
- The new Hipparcos reduction (van Leeuwen 2007)
- The extended Hipparcos compilation (Anderson & Francis 2012, XHIP)
- The ASCC-2.5 catalog (Kharchenko 2001) updated by Kharchenko et al. (2007)
- The PPMXL catalog (Roeser et al. 2010)
- The 2MASS catalog (Cutri et al. 2003)
- The Cepheid photometry obtained by Berdnikov et al. (2000); Berdnikov (2008)
- The McMaster Cepheid photometry and radial velocity data archive maintained by Doug Welch<sup>10</sup>
- The radial velocity data, see Sec. 3.2.3

<sup>8</sup> <http://www.aavso.org/vsx/>

<sup>9</sup> <http://www.astro.utoronto.ca/DDO/research/Cepheids/>

<sup>10</sup> <http://crocus.physics.mcmaster.ca/Cepheid/>

### 3.2.1 Cepheid Parallaxes

Parallax,  $\pi$ , is a key membership constraint, since cluster membership is virtually guaranteed if a Cepheid occupies the same space volume as a cluster. We combine parallax estimations from different sources, favoring *PLR-independent* determinations.

Parallax in [mas] is given preference over distance in [pc] here, since the uncertainty,  $\sigma_\pi$ , is normally distributed, in contrast to the error in distance. This is important, since the computation of likelihoods by Eq. 7 assumes Gaussian uncertainties.

We compile parallaxes in the following order of preference: i) Hipparcos absolute parallaxes and associated errors from the new reduction by van Leeuwen (2007, 5 Cepheids), so long as  $\sigma_\pi/\pi \leq 0.1$  and  $\pi > 0$ ; ii) HST-based parallaxes from Benedict et al. (2007, 8 Cepheids); iii) parallaxes computed from distances in Storm et al. (2011, 67 Cepheids) including their uncertainties; iv) parallaxes calculated from PLR-based distances (624 Cepheids), see below.

PLR-based parallaxes of fundamental-mode Cepheids are calculated from distances computed following Turner et al. (2010). Our choice of P-L relation was motivated mainly by the considerations that i) V-band magnitudes can be obtained for the largest number of Cepheids; ii) the above formulation is calibrated for the Galaxy using the most recent observational results, including the HST parallaxes by Benedict et al. (2007) and the cluster Cepheids from Turner (2010).

We thus calculate PLR distances as follows:

$$5 \log d = \langle m_V \rangle - \langle M_V \rangle - A_V + 5, \quad (9)$$

where  $\langle m_V \rangle$  is the apparent mean V-band magnitude, and the average absolute V-band magnitude,  $\langle M_V \rangle$ , is obtained from the pulsation period  $P$  via:

$$\langle M_V \rangle = -(1.304 \pm 0.065) - (2.786 \pm 0.075) \log P. \quad (10)$$

Eq. 10 is valid only for fundamental-mode pulsators, no distances were estimated for overtone pulsators. The total absorption,  $A_V$ , is defined as:

$$A_V = R_V \cdot E(B - V), \quad (11)$$

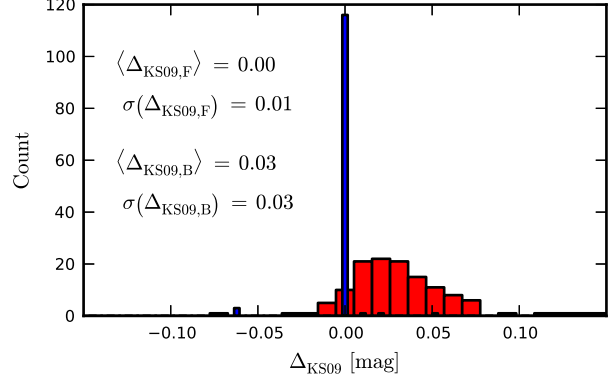
with  $R_V = 3.1$  the canonical ratio of total to selective extinction (reddening law) and  $E(B - V)$  the color excess of the object, cf. Sec. 3.2.1.2. The parallax is simply:

$$\pi = \frac{1000}{d} [\text{mas}] \quad (12)$$

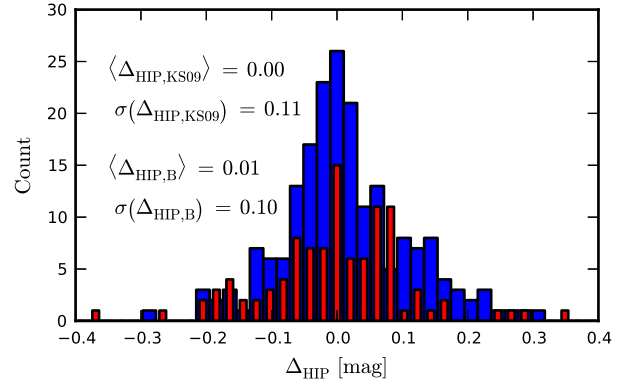
with  $d$  in [pc]. The parallax uncertainty,  $\sigma_\pi$ , is obtained considering the error budget on the distance,  $\sigma_d$ :

$$\sigma_\pi = \frac{1000}{d^2} \cdot \sigma_d [\text{mas}]. \quad (13)$$

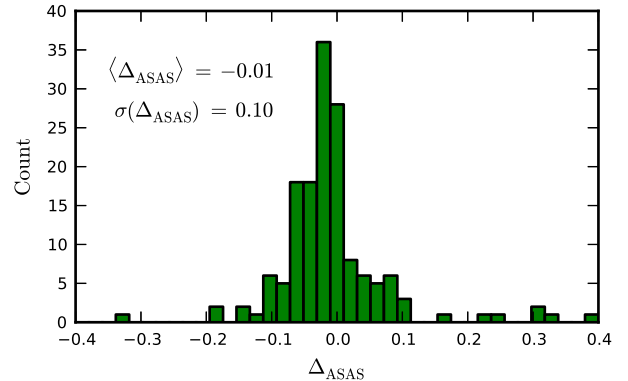
Thus, to estimate a Cepheid's parallax, knowledge of the PLR,  $P$ ,  $\langle m_V \rangle$ , and  $A_V$  is required. Periods are usually available in the GCVS or the VSX, whereas average magnitudes and color excesses of many of the newer Cepheid candidates are not available in the literature. However,  $E(B - V)$  can be estimated from combined NIR and optical data. The following paragraphs describe in detail how these quantities are compiled.



**Figure 7.** Histogram of differences in mean magnitudes relative to the KS09 values,  $\Delta_{\text{KS09}}$ . Blue slim bars show  $\Delta_{\text{KS09,F}}$  computed using Fernie values (123 Cepheids); Red broad bars show  $\Delta_{\text{KS09,B}}$  computed using data from Berdnikov et al. 2000 (127 Cepheids). Mean differences,  $\langle \Delta_{\text{KS09}} \rangle$ , and dispersions,  $\sigma(\Delta_{\text{KS09}})$ , given in [mag].



**Figure 8.** Histogram of differences in mean magnitudes relative to Hipparcos median V-band,  $\Delta_{\text{HIP}}$ . Red slim bars show  $\Delta_{\text{HIP,KS09}}$  computed using KS09 (104 Cepheids). Blue broad bars show  $\Delta_{\text{HIP,B}}$  computed using Berdnikov et al. 2000 (198 Cepheids). Mean differences,  $\langle \Delta_{\text{HIP}} \rangle$ , and dispersions,  $\sigma(\Delta_{\text{HIP}})$ , given in [mag].



**Figure 9.** Histogram of differences in mean V-band magnitude computed from ASAS light curves for 154 Cepheids relative to reference values from Fernie and KS09,  $\Delta_{\text{ASAS}}$ , cf. also Fig. 7. Mean difference,  $\langle \Delta_{\text{ASAS}} \rangle$ , and dispersion,  $\sigma(\Delta_{\text{ASAS}})$ , given in [mag].

**3.2.1.1 Mean Magnitude,  $\langle m_V \rangle$**  We compile mean V-band magnitudes,  $\langle m_V \rangle$ , from multiple references. Different methods of determining mean magnitudes exist, and the photometry employed is inhomogeneous, forcing us to adopt a zero-point for mean magnitudes compiled. In Fig. 7 we therefore compare mean magnitudes from Klagyivik & Szabados (2009, from hereon: KS09) with the Fernie et al. (1995) database’s magnitude-based means and the intensity-means from Berdnikov et al. (2000). For the Cepheids common to both studies, KS09 and the Fernie magnitudes show excellent agreement. We therefore adopt the following order of preference for compiling mean magnitudes.

First, we adopt  $\langle m_V \rangle$  values from KS09 with a fixed error budget of 0.03 mag, since the study carefully investigates amplitudes with a special focus on binarity.

Second, we adopt the magnitude-based means from the Fernie database with uncertainties calculated as the difference between intensity- and magnitude-mean magnitudes, with a minimum error of 0.03 mag.

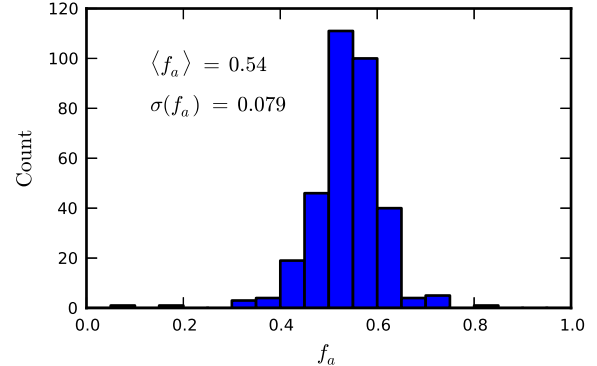
Third, we include Berdnikov et al. (2000) mean magnitudes.. As seen in Fig. 7, these  $\langle m_V \rangle$  values are systematically smaller (brighter) by approx. 0.03 mag than KS09. This discrepancy is most likely due to different ways of determining the mean. We remove this offset from the Berdnikov et al. (2000) values for internal consistency and adopt 0.03 mag as error budget for these values, identical to KS09.

Fourth, we employ median V-band magnitudes from the Hipparcos catalog (Perryman & ESA 1997, obtained via XHIP) for 8 Cepheids. The median V-band magnitudes derived from Hipparcos magnitudes can differ significantly from mean magnitudes listed in other references, cf. Fig. 8. Usually, this is due to contamination due to a nearby star within the instantaneous field of view. As error budget for the Hipparcos median V-magnitudes, we adopt the standard deviation computed from the red histogram in Fig. 8, i.e. 0.110 mag. We note that we could find no dependence on period or number of transits for this dispersion.

Fifth, we adopt average apparent V-band magnitudes that we determine from ASAS-3 light curves. To this end, we fit Fourier Series (same procedure as described for radial velocities in Sec. 3.2.3.3) to the phased light curves and use the constant term as the average,  $\langle m_V \rangle$ . Figure 9 shows a histogram of  $\Delta_{\text{ASAS}}$ , the differences between the computed ASAS-based  $\langle m_V \rangle$  and Fernie or KS09. We remove the offset of  $-0.01$  mag from the ASAS mean magnitudes and adopt the dispersion of 0.10 mag computed as the error budget.

The large dispersion,  $\sigma(\Delta_{\text{ASAS}})$ , in Fig. 9 probably originates from contamination due to nearby stars. To illustrate this, Figure 11 shows phase-folded ASAS-3 V-band light curves of two Cepheids, CY Car (left) and BMPup (right). Our mean magnitude agrees well with the literature value for CY Car. For BMPup however, a systematic difference of approximately 0.144 mag is evident, although the light curve appears to be clean otherwise. Inspection of a DSS images, however, reveals that contamination from a nearby companion is likely. Out of 154 Cepheids for which the ASAS light V-band curves were inspected, 20 differed by more than 0.1 mag from the reference value, and 28 agreed to within 0.01 mag.

If no mean magnitude is obtained from any of the



**Figure 10.** Fractional amplitudes,  $f_a$ , computed from Fernie mean, and GCVS maximum and minimum magnitudes, see Eq. 14.

above sources, we perform a (rough) estimate of  $\langle m_V \rangle$  based on the information provided in the GCVS and the VSX, using the magnitude at maximum brightness,  $\min_V$ , and the amplitude,  $\text{amp}_V$ , of the V-band light curve.  $\text{amp}_V$  is either provided directly by the catalogs, or calculated as the difference between minimum and maximum brightness,  $\text{amp}_V = \max_V - \min_V$ .

Since Cepheid light curves are skewed, their mean magnitudes do not necessarily lie at half the amplitude. We therefore estimate the typical fractional amplitude at mean brightness,  $\langle f_a \rangle$ , to compute  $\langle m_V \rangle = \min_V + \langle f_a \rangle \text{amp}_V$ . Figure 10 shows a histogram of  $f_a$  computed using mean magnitudes listed in the Fernie database,  $\langle m_{V,F} \rangle$ , and amplitudes,  $\text{amp}_V$ , from the catalogs. We find

$$\langle f_a \rangle \equiv \text{median} \left( \frac{\langle m_{V,F} \rangle - \max_V}{\min_V - \max_V} \right) = 0.54 \pm 0.079. \quad (14)$$

We derive an uncertainty on  $\langle m_V \rangle$  thus obtained using the uncertainty on  $\langle f_a \rangle$ , an estimated error on the amplitude, and a prescribed error on the magnitude at maximum brightness (we adopt 0.1 mag for 12th magnitude and below, and increase linearly to 0.5 mag at 20th magnitude). The resulting mean error on  $\langle m_V \rangle$  is 0.27 mag.

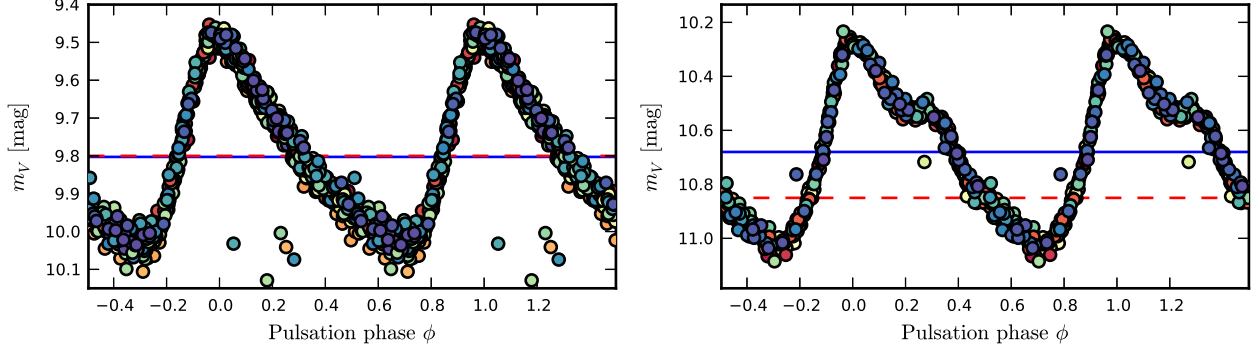
This estimation works reasonably well, although there exist obvious limitations, such as inhomogeneity of passbands, accuracy of the upper and lower limits, the applicability of the above ratio for a given Cepheid. Nevertheless, it does provide access to rough estimates of  $\langle m_V \rangle$  for Cepheids with little available information.

Given the many different ways in which average magnitudes were estimated, we keep track of the type of estimation to ensure traceability of any potential issues.

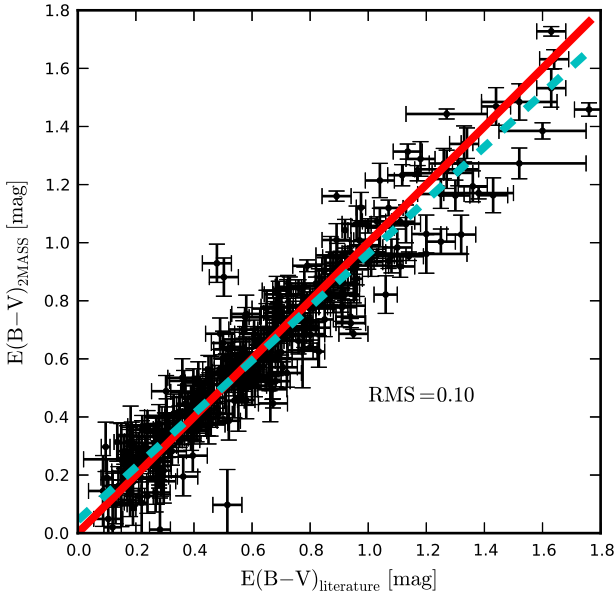
**3.2.1.2 Color excess,  $E(B - V)$**  The principal reference adopted for color excess is Kovtyukh et al. (2008) from which we also adopt stated uncertainties,  $\sigma(E(B - V))$ . Where no  $\sigma(E(B - V))$  were listed, we adopt an error budget of 0.05 mag, which is the mean of the given uncertainties.

For Cepheids without measurements in Kovtyukh et al. (2008), we adopt  $E(B - V)$  from the Fernie database and an error budget of 0.05 mag, unless the standard error from multiple reddening estimations was given.





**Figure 11.** Phase-folded ASAS-3 V-band light curves of CY Car (left) and BM Pup (right). Julian date of observation is color-coded, increasing from red to blue. Horizontal lines indicate reference average magnitude (red dashed, KS09) and constant term of the fitted Fourier Series (blue solid). For CY Car, the two are in excellent agreement. BM Pup has a bright neighbor that contaminates the aperture used to measure its flux, leading to an underestimated (too bright) mean magnitude.



**Figure 12.** Color excess from the literature against 2MASS-based estimate. Red solid line indicates the diagonal, cyan colored dashed line represents weighed least squares fit. RMS around either line: 0.10 mag.

Further  $E(B-V)$  values were adopted from Fouqué et al. (2007) together with the stated uncertainties.

Color excesses for other Cepheids are estimated using single-epoch 2MASS (Cutri et al. 2003) J-band magnitudes,  $m_J(\text{JD})$ , following Majaess et al. (2008). The method requires knowledge of the pulsation period,  $P$ , and the average J-band magnitude,  $\langle m_J \rangle$ . To obtain the latter, we employ their Eq. 5 that reads:

$$\langle m_J \rangle \simeq m_J(\text{JD}) - \left[ \frac{|m_{V(\phi_J)} - \max_V|}{\text{amp}_V} - 0.5 \right] \cdot 0.37 \text{amp}_V, \quad (15)$$

where  $\phi_J$  denotes the pulsation phase of the J-band measurement, and  $m_{V(\phi_J)}$  is the V-band magnitude at that phase.  $E(B-V)$  can then be estimated by

$$E(B-V) = -0.270 \log P + 0.415 (\langle m_V \rangle - \langle m_J \rangle) - 0.255. \quad (16)$$

Wherever possible,  $m_{V(\phi_J)}$  was obtained from the ASAS light curve. If this is impossible, we assume a sinusoidal light curve with the given mean magnitude and (semi-)amplitude.

Uncertainties or changes in pulsation period can significantly impact the phase calculated for the single-epoch 2MASS measurement,  $\phi_J$ . We therefore optimize Cepheid ephemerides for which ASAS data were available. To do so, we compute a grid (at fixed periods) of Fourier Series fits around the period provided in the ACVS and retain the solution with the minimum root mean square. Epochs are optimized by simply shifting the phase-folded curve. Figure 13 illustrates this step for the overtone Cepheid QZ Nor. We then take care to employ the most recently determined pulsation ephemerides available and estimate reddening uncertainties using error propagation for the quantities involved.

We note that this approach may be subject to multiple issues such as: i) the unknown shape of the light curve; ii) the applicability of Eq. 16; iii) period changes that impact  $m_{V(\phi_J)}$ ; iii) the approximate form of the relationship in Eq. 15. We therefore compare the 2MASS-based color excesses to the reference values, see Fig. 12, where the result of a weighted least squares fit is indicated by a dashed cyan line and does not differ much from the diagonal indicated by a solid red line. Despite considerable dispersion, the correspondence is clear and the results are promising (RMS of 0.1 mag).

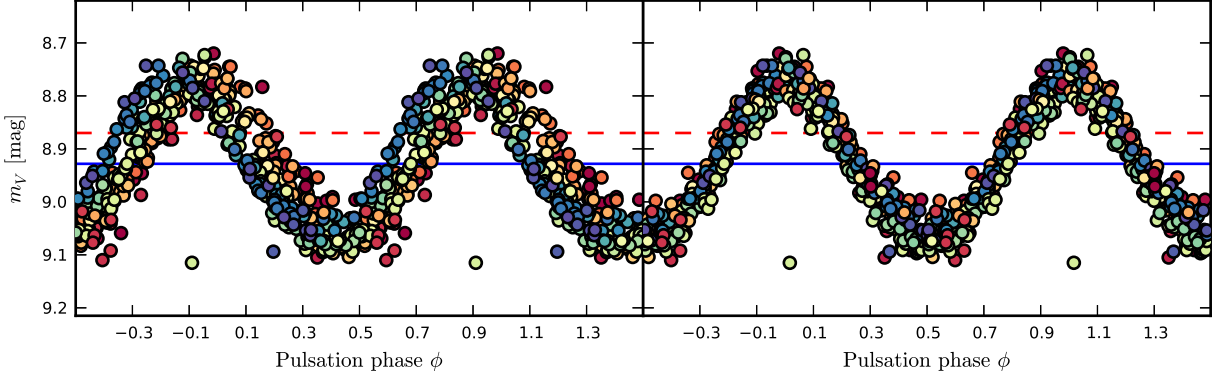
### 3.2.2 Proper Motions

Cepheid proper motions are taken from the following sources in order of preference:

- (i) Hipparcos proper motions from the new reduction by van Leeuwen (2007)
- (ii) The PPMXL catalog by Roeser et al. (2010)

### 3.2.3 Systemic Radial Velocities

**3.2.3.1 New Observations** In order to extend the number of Cepheids with known systemic radial velocities,



**Figure 13.** ASAS-3 V-band phase-folded light curve of QZ Nor. Color-code from red to blue indicates increasing Julian date of observation. *Left panel:* pulsation period and epoch of maximum light from the ACVS. *Right panel:* optimized period and epoch used.

$v_\gamma$ , we<sup>11</sup> carried out observations between November 2010 and July 2012 using the fiber-fed high-resolution echelle spectrographs *CORALIE* (Queloz et al. 2001, see also the instrumental upgrades described in Ségransan et al. 2010,  $R \sim 60000$ ) at the 1.2 m Euler telescope at La Silla, Chile, and *HERMES* (Raskin et al. 2011,  $R \sim 80000$ ) at the identically-built Mercator telescope on La Palma. In total, we observed 103 Cepheids with *CORALIE* and 63 with *HERMES*. 18 Cepheids were observed with both instruments, i.e. from both hemispheres. For 85 of these Cepheids, no radial velocity (RV) data are available in the literature.

Efficient reduction pipelines exist for both instruments that include pre- and overscan bias correction, cosmic removal, as well as flatfielding using Halogen lamps and background modelization. ThAr lamps are used for the wavelength calibration.

The RVs are computed via the cross-correlation technique described in Baranne et al. (1996). We use numerical masks designed for solar-like stars (optimized for spectral type G2) for all cross-correlations. Both instruments are very stable and yield very high precision RVs of  $\sim 10 \text{ ms}^{-1}$  (Queloz et al. 2000; Raskin et al. 2011). The measurement uncertainty is therefore not limited by the instrumental precision, but by line asymmetries due to pulsation. A detailed investigation of these effects is out of scope for this paper and will be presented in a future publication. The typical uncertainty on individual measurements is thus at the  $100 - 300 \text{ ms}^{-1}$  level, depending on the star and pulsation phase.

**3.2.3.2 Literature Data** In addition to the previously unpublished radial velocities described in Sec. 3.2.3.1, we employ literature data from many references to determine systemic velocities,  $v_\gamma$ , see Sec. 3.2.3.3. The addition of literature RVs extends the baseline of our otherwise relatively short (1.5 years) observing program, thereby enhancing our sensitivity to binarity. For binary Cepheids with published orbital solutions, we adopt the literature  $v_\gamma$ .

Aside from the systemic RVs in the Fernie database and KS09, we compile RV time series from the follow-

ing sources: Lloyd Evans (1980); Gieren (1981, 1985); Evans (1983); Coulson et al. (1985); Coulson & Caldwell (1985); Barnes et al. (1987, 1988); Gieren et al. (1989); Wilson et al. (1989); Metzger et al. (1991, 1992, 1998); Gorynya et al. (1992); Evans & Welch (1993); Pont et al. (1994, 1997); Bersier et al. (1994); Kiss (1998); Imbert (1999); Storm et al. (2004); Barnes et al. (2005); Petterson et al. (2005); Baranowski et al. (2009). The data for most of these sources dated earlier than 1986 are extracted from the McMaster Cepheids database<sup>12</sup>. Newer data are obtained through VizieR<sup>13</sup>.

**3.2.3.3 Systemic Radial Velocities,  $v_\gamma$**  The systemic radial velocity,  $v_\gamma$ , is obtained by fitting a Fourier series to the RVs. We use pulsation period,  $P$ , as a fixed parameter, since it is known for all Cepheids we observed.

The basic analytical form applied was a Fourier Series with  $n$  harmonics and phase  $\phi$  is:

$$FS_n = v_\gamma + \sum_{n=1,2,3,\dots} a_n \sin(2n\pi\phi) + b_n \cos(2n\pi\phi) \quad (17)$$

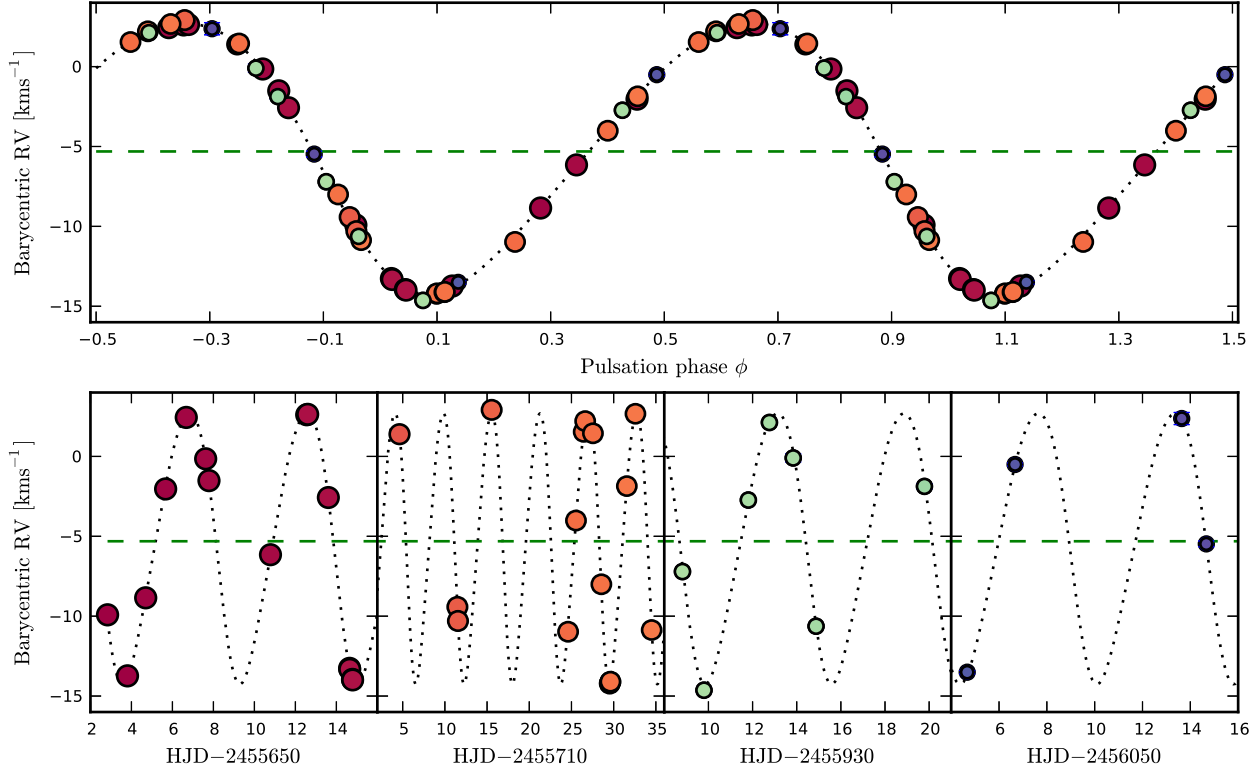
Since the number of data points available varies for each star, we do not fix the number of harmonics in this fit. Instead, we iteratively increase the degree of the Fourier series until an F-test indicates an overly complex representation, i.e. when spurious fit improvement is more likely than 0.27 per cent. For some stars, we therefore use only a simple sine function, whereas stars with many measurements are fitted using up to five harmonics. We show two examples of newly observed RV curves in Figs. 14 and 15. A full description and publication of the new radial velocity data will follow in the near future.

We adopt a fixed error budget of  $3 \text{ km s}^{-1}$  on  $v_\gamma$ . Although this may overestimate the uncertainties for some very good cases, it is intended to account for a range of systematic errors, such as unseen binarity, instrumental zero point differences and insufficient phase coverage. We are confident that this error budget is sufficiently large to prevent the exclusion of good member candidates, while being suffi-

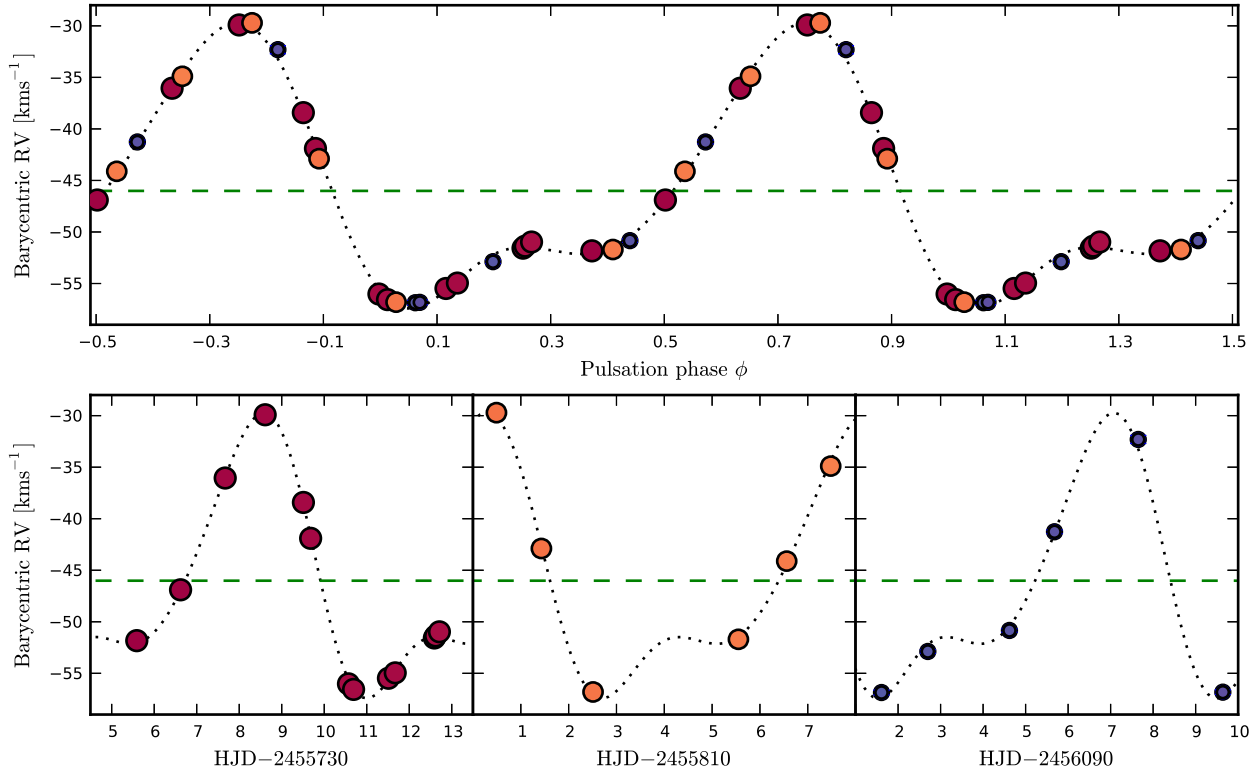
<sup>11</sup> The authors are grateful to the observers who contributed in this effort; names are given in the acknowledgments.

<sup>12</sup> maintained by Doug Welch, available at <http://crocus.physics.mcmaster.ca/Cepheid/>

<sup>13</sup> <http://vizier.u-strasbg.fr/viz-bin/VizieR>



**Figure 14.** RV data recently published in Szabados et al. 2012 for first-overtone pulsator GH Car, obtained in southern hemisphere with *CORALIE* as part of our program. Color-code (red to blue) and size-code (larger to smaller) indicate increasing Julian Date of observation. Fit of triple-harmonic Fourier Series indicated by black dashed line.  $v_\gamma$  indicated by green dashed horizontal line. *Top panel:* phase-folded, *Bottom panel:* non-folded, showing zooms of the four observing runs during which data were obtained.



**Figure 15.** New RV data for fundamental-mode Cepheid V2340 Cyg, obtained in northern hemisphere with *HERMES*. Color-code (red to blue) and size-code (larger to smaller) indicate increasing Julian Date of observation. Fit of triple-harmonic Fourier Series indicated by black dashed line.  $v_\gamma$  indicated by green dashed horizontal line. *Top panel:* phase-folded, *Bottom panel:* non-folded, showing zooms of the three observing runs during which data were obtained.

ciently small to provide a stringent constraint. If insufficient data points render the Fourier fit unsatisfactory, we determine a rough estimate of  $v_\gamma$  and its error budget by eye.

### 3.2.4 Iron Abundance

We rely mostly on iron abundances by Luck & Lambert (2011) and complement these with the compilation in KS09 that made an effort to homogenize measurements from many references, notably Andrievsky and collaborators (e.g. Luck & Andrievsky 2004). Unfortunately, the adopted standard value of solar iron abundance can vary among references, possibly introducing systematic offsets between different authors' studies. Furthermore, an estimation of the iron abundance in a Cepheid is more complex than in a non-pulsating star, since the stellar parameters (e.g. temperature and turbulence) vary during the pulsation cycle, and since the atmosphere is not static. We therefore adopt generous error budgets, 0.1 dex for the values from Luck & Lambert (2011), and 0.15 dex in [Fe/H] for the compiled list in KS09.

### 3.2.5 Age

Cepheid ages can be calculated for first overtone and fundamental mode pulsators using the period-age (PA) relations given in Bono et al. (2005). For fundamental mode Cepheids, we use  $\log t = (8.31 \pm 0.08) - (0.67 \pm 0.01) \log P$ . For overtone pulsators, the relation used is  $\log t = (8.08 \pm 0.04) - (0.39 \pm 0.04) \log P$ . Age error budgets are calculated from the uncertainties stated for slope and intercept.

## 4 RESULTS

This section presents the results from our census. We start with the *bona-fide* CCs (Sec. 4.1), followed by strong candidates that lie far from cluster center and therefore have a low prior (Sec. 4.1.2). For brevity of the main body, we defer presentation of inconclusive and rejected Combos that have been previously discussed in the literature, as well as cases of fore- and background contamination, to appendix A. A full table that includes all data used is available online through the CDS.

Each subsection starts with Combos previously discussed in the literature. We take care to include relevant references. However, in a field with this much history, it is almost inevitable that some works were overlooked.

As a sanity check of our results, we compare them to known CC candidates from the literature. The main references used for this benchmarking are the reviews by Feast (1999) and Turner (2010, from hereon: T10). However, there are further Cepheids whose cluster membership was considered in the literature that will be mentioned where appropriate. As mentioned in Sec. 3.1, 12 of the host clusters listed in T10 are not present in our cluster sample. For 8 of these, the Cepheids are cross-matched and investigated for membership in different clusters. However, none of these host cluster 'alternatives' was found to be a good match. Among these, SW Vel has previously been investigated for membership in NGC 2660 by Turner et al. (1993) who found the same null result. For the Cepheids X Cyg,  $\zeta$  Gem, V367 Sct,

**Table 1.** Membership probabilities of *bona-fide* cluster Cepheids. Important references: a: Irwin 1955, b: Kholopov 1956, c: Feast 1957, d: Sandage 1958, e: Eggen 1980, f: Turner 1982, g: Walker 1985, h: Turner 1986, i: Turner et al. 1992, j: Matthews et al. 1995 and references therein, k: Turner et al. 1997, l: Hoyle et al. 2003, m: An et al. 2007, n: Turner 2010, o: first discussed in this work. Column Data lists constraints employed 1: separation (relative position, cluster radius), 2:  $\pi$ , 3: RV, 4:  $\mu_\alpha^*$ , 5:  $\mu_\delta$ , 6: [Fe/H], 7: age.  $P(A|B) = P(A) \cdot P(B|A)$ , see Sec. 2

Cepheid	Cluster	Ref	Data	$P(A B)$	$P(B A)$
SU Cyg	Turner 9	k,n	1-5,7	0.946	0.946
U Sgr	IC 4725	a,b,c,m	all	0.983	0.983
DL Cas	NGC 129	b,i	1-5,7	0.928	0.928
CF Cas	NGC 7790	d,j	1-5,7	0.752	0.979
S Nor	NGC 6087	a,b,c,h	all	0.710	0.710
V340 Nor	NGC 6067	g,m,n,l	all	0.605	0.605
CE CasA	NGC 7790	d,j	1-5,7	0.540	0.977
CE CasB	NGC 7790	d,j	1-5,7	0.522	0.963
QZ Nor	NGC 6067	e,g	all	0.027	0.952
V Cen	NGC 5662	f,m,n	1-5,7	0.005	0.923
V379 Cas	NGC 129	o	1,3,4,5,7	0.000	0.923
SX Car	ASCC 61	o	1,2,4,5,7	0.001	0.922
S Mus	ASCC 69	o	1-5,7	0.004	0.879

and  $\alpha$  UMi (Polaris), no candidate host clusters were identified within our cluster sample at separations inferior to 5 limiting radii.

In Sec. 4.2, we employ our *bona-fide* CC sample in a calibration of the Cepheid period-luminosity relation.

### 4.1 Bona-fide Cluster Cepheids

Table 1 lists the *bona-fide* CCs recovered by our analysis in descending order of membership probability,  $P(A|B)$ . Column *Ref* contains the first and most relevant references discuss membership for a particular Combo, and column *Data* provides the constraints employed. The last two columns contain membership probability and likelihood.

The horizontal line in Tab. 1 separates the clear from the not-quite-so-clear CCs, the difference being that the latter are situated at larger separations from the cluster, i.e. have low priors. Among these, our work suggests the following new *bona-fide* CCs: S Mus in ASCC 69, SX Car in ASCC 61, and V379 Cas in NGC 129; see Sec. 4.1.2 for details.

#### 4.1.1 Bona-fide CCs with high $P(A|B)$

All of the CCs in the top part of Tab. 1 are well-established in the literature and are well constrained by the available information. They have very high likelihoods and high priors, i.e. they are rather obvious members. We therefore refer to the original references listed for more information, as well as to the data table supplied in electronic form<sup>14</sup>.

The case of S Nor in NGC 6087, however, deserves some closer inspection, since significantly different cluster RVs were found in the literature.

We note that a large fraction of the Cepheids in Tab. 1

<sup>14</sup> [Link / catalog number / ?](#)

are also binaries, see the binary Cepheid database<sup>15</sup> by L. Szabados, and references therein.

**4.1.1.1 S Nor and NGC 6087** S Nor’s membership in NGC 6087 was among the first to ever be suggested (Irwin 1955) and confirmed using radial velocities (Feast 1957), as well as the detailed study by Turner (1986) based on reddening and distance.

The values for the cluster’s mean RV differ greatly between D02 ( $6 \text{ km s}^{-1}$ ), K05 ( $-9 \text{ km s}^{-1}$ ), and (Feast 1957,  $2.0 \text{ km s}^{-1}$ ), which is important considering the Cepheid’s  $v_\gamma = 2.53 \text{ km s}^{-1}$  (Groenewegen 2008). We note that the value adopted by D02 is measured on the Cepheid itself (Mermilliod et al. 2008), although without taking orbital motion into account, and is therefore not suitable as a membership constraint. Since Feast (1957) investigated the largest number of stars and specifically targeted this cluster, we trust that their  $2.0 \text{ km s}^{-1}$  is the best available estimate for the mean velocity of the cluster.

Small differences exist between the cluster and Cepheid values for reddening, metallicity, and age. However, these differences barely exceed the combined uncertainties and may not be significant.

#### 4.1.2 Bona-fide CCs with Low Priors

The following Combos require a short discussion, since we obtain high likelihoods and low priors, resulting in low membership probabilities. We note that IRSB parallaxes from Storm et al. (2011) were employed for all but one of the Cepheids discussed here. For SX Car, we adopt a fundamental-mode PLR parallax using the E(B-V) value from the Fernie database.

**4.1.2.1 QZ Nor and NGC 6067** The overtone pulsator QZ Nor is a second Cepheid member to the rich open cluster NGC 6067 that also contains V340 Nor. Its cluster membership was first considered by Eggen (1983). Our likelihood of 95 per cent is based on all membership constraints. The radius determination for the cluster is rather difficult, since it is located in the Norma cloud, cf. the atlas page in K05. The low prior may therefore be misleading. Alarming, the IRSB parallaxes of QZ Nor ( $0.74 \pm 0.07 \text{ mas}$ ) and V340 Nor ( $0.58 \pm 0.09 \text{ mas}$ ) are nearly consistent with each other. However, their difference barely exceeds the combined uncertainty, and NGC 6067’s parallax ( $0.71 \text{ mas}$ ) lies between the two, consistent with both Cepheids within the error budgets adopted. It is possible that differential reddening could be of importance in resolving this conundrum (higher for V340 Nor which is closer to cluster center). Furthermore, iron abundance and radial velocity of both Cepheids are in very good agreement. Proper motion does not exceed the uncertainties and is therefore of limited use as a constraint. We conclude that the evidence combined in the likelihood is in favor of membership, which leads us to consider QZ Nor a *bona-fide* member of NGC 6067.

**4.1.2.2 V Cen and NGC 5662** V Cen’s available constraints (all but [Fe/H]) are in impressively close agreement with those of NGC 5662, yielding a very high likelihood (92 per cent). As for NGC 6067, the determination of the limiting radius is complicated by a very strongly populated area. Furthermore, the common proper motion clearly shows the cluster to be co-moving with the Cepheid. The slightly higher proper motion of the Cepheid does not appear to deviate significantly from that of the cluster. We therefore see no reason to doubt V Cen’s membership in NGC 5662.

**4.1.2.3 S Mus and ASCC 69** The membership constraints considered (all but [Fe/H]) for the binary Cepheid S Mus (Peterson et al. 2005) and ASCC 69 yield a very high likelihood of 86 per cent. ASCC 69 is a sparsely populated cluster for which K05 list merely twelve  $1\text{-}\sigma$  members. The radius of the cluster is therefore rather imprecise. Furthermore, radial velocities are only of limited value as membership constraints, since the average cluster RV is based on only 2 stars. However, proper motion clearly indicates that cluster and Cepheid are co-moving. We tentatively accept S Mus as a cluster member and stress the need for a detailed study of its candidate host cluster ASCC 69.

**4.1.2.4 SX Car and ASCC 61** The 4.86 d Cepheid SX Car is seen to be co-moving with ASCC 61 in proper motion. Unfortunately, no average cluster RV is known. Parallax and age, however, agree very well between cluster and Cepheid, lending support to the hypothesis of membership with  $P(B|A) = 95$  per cent. An in-depth analysis of the cluster, including its mean RV, is of the essence, since the cluster is located in a crowded field (K05). We tentatively consider SX Car as a member of ASCC 61.

**4.1.2.5 V379 Cas and NGC 129** This high likelihood pair at large separation ( $44'$  or  $3.4 \text{ pc}$  at the estimated distance to NGC 129) has nearly vanishing proper motion, while the both RVs are in excellent agreement and the ages are consistent. We obtain a likelihood of 92 per cent. No parallax is computed, since V379 Cas is an overtone pulsator. Since NGC 129 has another known member, DL Cas, we can compare the pulsational ages of the two Cepheids and find both to be consistent within the uncertainties ( $7.70 \pm 0.08$  for DL Cas and  $7.83 \pm 0.07$  for V379 Cas). Furthermore, the iron abundances of both Cepheids are close, as are their RVs (to within less than  $1 \text{ km s}^{-1}$ ). We further note that V379 Cas and DL Cas have similar reddening values with that of V379 Cas slightly ( $0.1 \text{ mag}$ ) higher (Kovtyukh et al. 2008, both). We therefore tentatively consider V379 Cas a member of NGC 129’s Halo, pending a better distance estimate and additional membership constraints. NGC 129 is thus particularly interesting, containing both a fundamental-mode and an overtone pulsator, just as NGC 6067.

#### 4.1.3 Chance Alignment

Among our *bona-fide* CCs, 5 are found to lie within  $r_c$ , and 4 at  $r_c < R < r_l$ . Within the original cross-match, 21 Combos are matched within  $r_c$  and 194 within  $r_c < R < r_l$ . Thus we can estimate the rate of chance alignment within the core radius to be about four to one (21/5). At separations inferior to  $r_l$ , this ratio increases to approximately 24 to 1 (215/9).

<sup>15</sup> <http://www.konkoly.hu/CEP/nagytab3.html>



**Table 2.** Membership probabilities calculated for inconclusive Combos. Constraints by numbers: 1=separation (i.e. position & cluster radius), 2= $\pi$ , 3=RV, 4= $\mu_\alpha^*$ , 5= $\mu_\delta$ , 6=[Fe/H], 7=age.  $P(A|B) = P(A) \cdot P(B|A)$ , see Sec. 2. AH 1 is the abbreviation for cluster Aveni-Hunter 1.

Cepheid	Cluster	Constraints	$P(A B)$	$P(B A)$
BD+47 4225	AH 1	1,4,5,7	0.408	0.947
VZ CMa	Trumpler 18	1,4-7	0.000	0.770
Y Sgr	IC 4725 (M 25)	all	0.000	0.702
GH Car	Trumpler 18	1,3,4,5,7	0.104	0.574
BB Sgr	Collinder 394	1-5,7	0.066	0.468
Y Car	ASCC 60	1-5,7	0.337	0.337
CG Cas	Berkeley 58	1,2,4,5,7	0.013	0.042
CV Mon	vdBergh 1	1-5,7	0.037	0.037
WZ Car	ASCC 63	1-5, 7	0.000	0.000

**Table 3.** Membership probabilities calculated for Combos inconsistent with membership. Italic cluster names indicate ‘alternative’ hosts considered. Constraints by numbers: 1=separation (i.e. position & cluster radius), 2= $\pi$ , 3=RV, 4= $\mu_{RA}$ , 5= $\mu_{DE}$ , 6=[Fe/H], 7=age.  $P(A|B) = P(A) \cdot P(B|A)$ , see Sec. 2

Cepheid	Cluster	Constraints	$P(A B)$	$P(B A)$
CS Vel	Ruprecht 79	1-5,7	0.148	0.148
V1726 Cyg	Platais 1	1,3,4,5,7	0.084	0.120
V442 Car	NGC 3496	1,4,5	0.013	0.035
TV CMa	NGC 2345	1-5,7	0.000	0.026
WZ Sgr	Sgr OB7	1,2,3,7	0.000	0.012
SZ Tau	NGC 1647	1-5,7	0.000	0.010
UY Per	Czernik 8	1,2,4,5,7	0.000	0.001
SU Cas	<i>FSR 0581</i>	1,2,4,5,7	0.000	0.000
V1334 Cyg	<i>ASCC 113</i>	1,3-7	0.000	0.000
TW Nor	<i>Nor OB5</i>	1,2,3,7	0.000	0.000
RU Sct	<i>Dolidze 34</i>	1	0.000	0.000
SW Vel	<i>NGC 2660</i>	all	0.000	0.000
S Vul	<i>Turner 9</i>	1-5,7	0.000	0.000
SV Vul	<i>Turner 9</i>	1-5,7	0.000	0.000

The remaining 4 *bona-fide* CCs are not included here, since they were found to lie outside the cluster radii defined. 3 of them are located within two  $r_1$ . V379 Cas lies at  $4.8 r_1$  from its host cluster’s center coordinates.

#### 4.1.4 Inconclusive Combos and Non-members

Combos for which our results are inconclusive are listed in Tab. 2. The Cepheids GSC 03642-02459 (BD+47 4225), Y Car, and WZ Car have thus far not been considered for cluster membership in the literature.

Two kinds of Combos inconsistent with membership are listed in Tab. 3. Italic cluster names identify Combos for which only ‘alternative’ host clusters were considered, i.e. ones whose literature host clusters were not included in the cluster sample considered. The remaining entries in this table are Combos considered inconsistent with membership, but have been previously considered for membership in the literature.

More details on the Combos listed in Tabs. 2 and 3 are provided in the appendix.

## 4.2 The Galactic Cepheid PLR revisited

Let us now employ our *bona-fide* CC sample in a calibration of the Galactic Cepheid PLR. It represents an ideal sample to this end, due to the high confidence we can have in cluster membership.

A Cepheid’s absolute V-band magnitude,  $M_V$ , is determined using the true distance modulus of the cluster,  $(V_0 - M_V)_{\text{Cl}}$ , the Cepheid’s mean magnitude,  $\langle m_V \rangle$ , the ratio of total-to-selective extinction towards the cluster,  $R_{V,\text{Cl}}$ , and the Cepheid’s color excess,  $E(B - V)$ , as

$$M_V = \langle m_V \rangle - (V_0 - M_V)_{\text{Cl}} - R_{V,\text{Cl}} E(B - V), \quad (18)$$

where quantities refer to the Cepheid, unless subscripted by ‘Cl’.

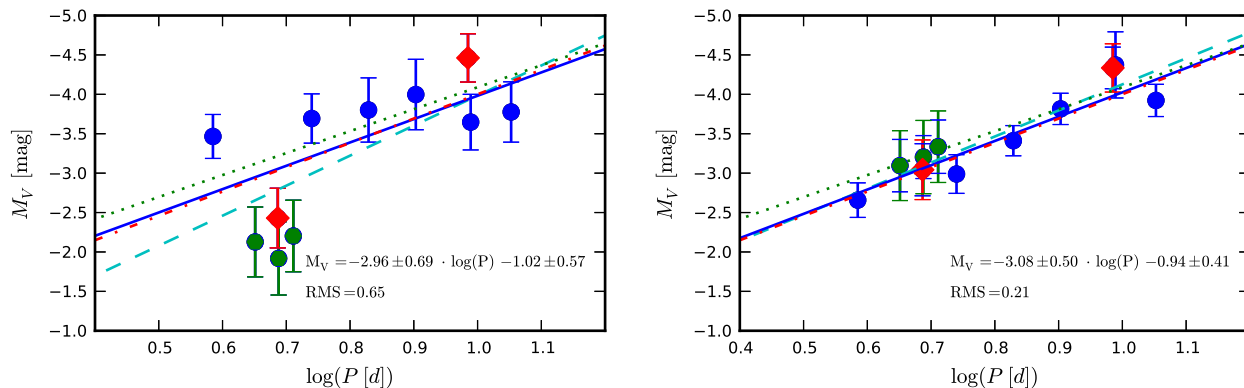
As described in Sec. 3, we compile cluster data from the references that yield the largest limiting radii. Obviously, this preference (bias) does not necessarily select the ‘best’ data available. For instance, line-of-sight dependencies of reddening are not taken into account by K05 and B11, and the ratio of total-to-selective absorption is fixed at the canonical value of 3.1. However, there may be a benefit in using the cluster data from these large catalogs, since the data are relatively homogeneous. In the following, we refer to the set of data compiled for our membership analysis as the ‘standard’ set.

In an attempt to improve accuracy, we compile data from detailed studies of the host clusters. Accurate true distance moduli based on ZAMS-fitting for many clusters can be found in T10, for instance. The values given are partially based on published literature and partially re-analyzed using 2MASS and unpublished UBV data (Turner 2012, priv. communication). For NGC 7790, we adopt the detailed work by Gupta et al. (2000), and for the two ASCC clusters, we use the de-reddened values by Kharchenko et al. (2005a). For reddening, An et al. (2007) provide a convenient way to calculate  $R_V$  for most lines of sight in our sample as  $R_V = R_{V,0} + 0.22(B - V)_0$ , with  $R_{V,0}$  tabulated for the cluster and taking into account the intrinsic color of the Cepheids. For the two ASCC clusters discovered by Kharchenko et al. (2005a), we employ  $R_V = 3.1$ , since no other estimate is available. For Turner 9 we use  $R_V = 2.94$  as in Turner et al. (1997). The Cepheid mean V-magnitudes and  $E(B - V)$  were compiled as described in Sec. 3; the mean magnitudes were taken typically from KS09, and the  $E(B - V)$  either from Kovtyukh et al. (2008) or from the Fernie database. The values employed are listed in Tab. 4, and we refer to this data set as the ‘optimal’ one.

Figure 16 shows the PLR fits to both the ‘standard’ (left panel) and the ‘optimal’ (right panel) data sets. In both figures, four straight lines indicate the PLR calibrations by Tammann et al. (2003, red dash-dotted), Turner et al. (2010, green dotted, lower zero-point), as well as our non-weighted (dashed cyan), and weighted (solid blue) least-squares fits to the data. The large scatter (RMS=0.65) in the ‘standard’ data set is striking. Contrastingly, the ‘optimal’ set appears worthy of its name, with an RMS of 0.21 mag. We determine the uncertainties on the ‘optimal’ set by linear regression and obtain:

$$M_V = -(3.08 \pm 0.50) \log P - (0.94 \pm 0.42). \quad (19)$$

Albeit by chance and with much larger uncertainty,



**Figure 16.** Cepheid PLRs fitted to the ‘standard’ (left panel) and ‘optimal’ (right panel) data sets. Solid circles indicate previously known CCs, green solid circles highlight the three Cepheids in NGC 7790. Two solid red diamonds indicate the two new *bona-fide* fundamental-mode CCs SMus and SX Car. The green dotted line shows the PLR by Benedict et al. 2007, and the red dash-dotted line represents Tammann et al. 2003. Blue solid and cyan dashed lines indicate weighted and non-weighted least-squares fits. An accurate PLR calibration critically depends on accurate distance estimates from detailed studies that include line-of-sight-variations of extinction.

Cepheid	Cluster	$\langle m_V \rangle$	$E(B - V)$	$(V_0 - M_V)$	$R_{V,Cl}$	$\log P$	$M_V$
U Sgr	IC 4725	6.72	$0.40 \pm 0.02$	$8.81 \pm 0.10$	$3.32 \pm 0.22$	0.829	$-3.41 \pm 0.19$
CF Cas	NGC 7790	11.15	$0.53 \pm 0.03$	$12.60 \pm 0.15$	$3.33 \pm 0.27$	0.688	$-3.20 \pm 0.27$
CE CasB	NGC 7790	10.94	$0.50 \pm 0.05$	$12.60 \pm 0.15$	$3.33 \pm 0.27$	0.711	$-3.34 \pm 0.34$
CE CasA	NGC 7790	11.09	$0.48 \pm 0.05$	$12.60 \pm 0.15$	$3.31 \pm 0.27$	0.651	$-3.10 \pm 0.33$
DL Cas	NGC 129	8.98	$0.49 \pm 0.02$	$11.11 \pm 0.02$	$3.46 \pm 0.23$	0.903	$-3.82 \pm 0.20$
SU Cyg	Turner 9	6.89	$0.07 \pm 0.02$	$9.33 \pm 0.20$	$2.94 \pm 0.39$	0.585	$-2.66 \pm 0.22$
V Cen	NGC 5662	6.87	$0.17 \pm 0.05$	$9.28 \pm 0.05$	$3.47 \pm 0.39$	0.740	$-2.90 \pm 0.25$
S Nor	NGC 6087	6.41	$0.27 \pm 0.05$	$9.78 \pm 0.03$	$3.74 \pm 0.86$	0.989	$-4.37 \pm 0.42$
V340 Nor	NGC 6067	8.41	$0.34 \pm 0.01$	$11.19 \pm 0.15$	$3.36 \pm 0.30$	1.053	$-3.92 \pm 0.21$
S Mus	ASCC 69	6.13	$0.15 \pm 0.03$	$10.00 \pm 0.20$	$3.10 \pm 0.51$	0.985	$-4.34 \pm 0.27$
SX Car	ASCC 61	9.12	$0.33 \pm 0.03$	$11.14 \pm 0.20$	$3.10 \pm 0.51$	0.687	$-3.04 \pm 0.33$

**Table 4.** Parameters adopted for the ‘optimal’ set used in Eq. 18 and the right panel of Fig. 16. True distance moduli of clusters,  $(V_0 - M_V)$ , and absorption-relevant parameters were adopted according to the criteria specified in the text. We adopt 0.04 mag as the uncertainty on  $\langle m_V \rangle$ .

this is nearly identical to the result of  $M_V = -(3.087 \pm 0.085) \log(P) - (0.914 \pm 0.098)$  obtained by Tammann et al. (2003). Owing to the large uncertainties of the fit, our solution is in agreement with most other zero-point-calibrated PLRs such as the ones in Fouqué et al. (2007); Turner et al. (2010). We do note, however, that the solution by Benedict et al. (2007) lies rather far from our calibration in both slope and intercept.

We note that no correction for binarity was applied in this calibration and that all Cepheids were treated as single stars, which many of them are not. The calibration is mainly limited by the accuracy of the cluster data, so that a correction for binarity is not likely to impact our result significantly.

## 5 DISCUSSION

### 5.1 Constraining Power and Limitations of Membership Constraints

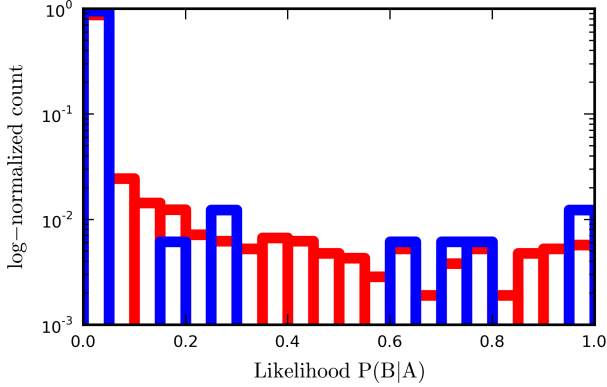
#### 5.1.1 The Prior $P(A)$

The form of the prior was motivated by radial density profiles of star clusters, see Sec. 2.1. However, the degree with

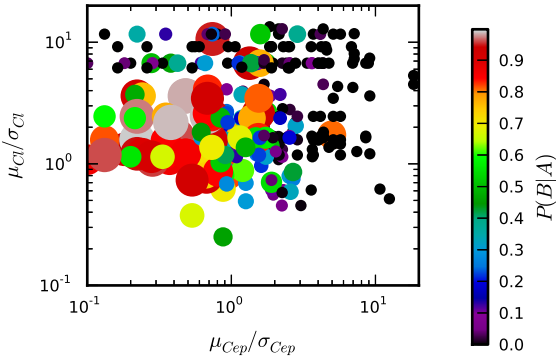
which the distribution of stars in the cluster is known varies greatly between clusters, and deviations from circular cluster shapes were ignored for internal consistency. Furthermore, crowding, great distances ( $\sim$  kpc) to host cluster candidates, differential reddening, sparsity, etc. all conspire to complicate the definition of cluster radii. It is thus not surprising that the prior does not perform extremely well as a membership constraint if taken at face value. Nevertheless, it does help to separate the interesting cases from the majority of null matches, since it reduces the question of proximity on the projected sky to a single number that contains information on the density of the cluster, since both cluster radii are used in its definition.

#### 5.1.2 The Likelihood $P(B|A)$

Intuitively, the greatest set of constraints used provides the tightest constraints on membership for any cluster-Cepheid combination (Combo). Figure 17 illustrates this. It shows a logarithmic normalized histogram of likelihoods for two cases: more than 3 constraints used to calculate  $P(B|A)$  (red); all constraints used to calculate  $P(B|A)$  (blue). However, the constraining power of a given set of constraints is not merely a function of its size. Here we discuss the mem-



**Figure 17.** Log-normalized histogram of likelihoods. Red: Combinations with 3 or more parameters used for  $P(B|A)$ . Blue: Combinations with all membership constraints. The more membership constraints are employed, the more separated are high and low likelihoods, i.e. the better constraint is membership.



**Figure 18.** Likelihoods computed as a function of proper motion. Abscissa: proper motion of Cepheid divided by the squared-summed uncertainties, i.e.  $\mu = \sqrt{(\mu_\alpha^*)^2 + \mu_\delta^2}$ ;  $\sigma = \sqrt{\sigma(\mu_\alpha^*)^2 + \sigma(\mu_\delta)^2}$ . Ordinate: same for Cluster.

bership constraining power of the different constraints used to calculate likelihoods.

**Proper motions** can be very effective at ruling out membership if the motion clearly exceeds the uncertainties on the measurement. This is the case for a cluster cross-matched with a background Cepheid, for instance. However, for a significant fraction of Combos, the proper motion vector’s magnitude was smaller than the uncertainty of the measurement, thus effectively not constraining membership, see Fig 18. If the magnitude surpasses the uncertainties by at least a factor of 3, proper motion serves as a reliable constraint. For the majority of Combos that fulfill this criterion, membership tends to be excluded.

**Distance** is a potentially very strong membership constraint, since intuitively, a Cepheid that occupies the same space volume as a cluster should be a member. However, cluster distances can be subject to large systematic uncertainties due to parameter degeneracy (distance, age, reddening), model-dependence (rotation, etc.), or previous distance estimates to, e.g., the Pleiades. Furthermore, implicit assumptions on cluster membership can significantly impact

the distance determined, especially for relatively young clusters that harbor few stars around the Main-sequence turn-off and few or no red giants. Very detailed studies of open clusters, and in particular of the line-of-sight extinction, are required for improvement in this domain, as is shown in Sec. 4.2, or demonstrated by the discussion of CV Mon’s membership in van den Bergh 1 (cf. Sec. A1.0.3).

**Radial velocities** have the potential to provide very tight membership constraints, since the RV dispersion within open clusters can be significantly below  $1 \text{ km s}^{-1}$  (Lovis & Mayor 2007), approaching the measurement precision on  $v_\gamma$  for non-binary Cepheids. However, estimates of average cluster RVs are usually based on only a few stars, see Fig. 5. In fact, approximately half of all clusters with ‘known’ average RVs are based on measurements of two stars or less, and strong selection effects (e.g. toward late-type stars) can severely impact the estimate. Since Cepheids are very bright and of late spectral type, it is rather likely that a cluster RV is in part based on measurements that include a Cepheid. We therefore highlight the need to observe more radial velocities of upper main sequence stars in clusters in order to ensure the most accurate estimates of average cluster RV. Alternatively, larger telescopes may observe the much fainter lower main sequence. The Gaia-ESO public spectroscopic survey Gilmore et al. (2012) will soon provide precise RVs for a large number of clusters and thus can improve the reliability of a future study similar to the present one.

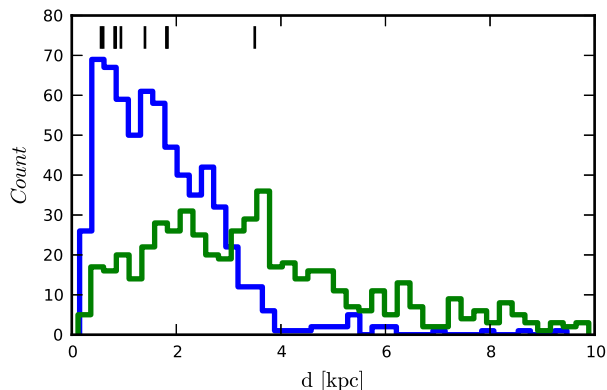
The **iron abundances** compiled here were, arguably, of limited use as membership constraints, since data inhomogeneity, the limited number of cluster stars used for determining the cluster average, and differences in the solar reference values used are at the same order of magnitude as the range of iron abundances found in the sample considered (that lies within the young metal-rich Galactic disk). For a few cases, however, the iron abundance did further strengthen the interpretation of excluded membership.

**Age** as a membership constraint quantifies valid evolutionary considerations that are established empirically. It is a particularly useful membership constraint, since it is readily available for clusters as well as for Cepheids (from period-age relations), especially when few other constraints are available. However, ages for both kinds of objects are subject to model-dependence, and the accuracy of the values inferred is difficult to quantify.

All of the above quantities have their own peculiarities, and thus no single one can be named the ‘best’ membership constraint. Instead, the greatest constraining power resides in the combination of all the data, as is seen in the *bona-fide* CCs identified in this work, as well as in Fig. 17.

### 5.1.3 Membership Probabilities $P(A|B)$

We computed  $P(A|B)$  simply as the product of the prior and the likelihood, leaving out the normalization term  $P(B)$  that would in principle be required, see Eq. 1. However, in order to make full use of the  $P(A|B)$  values computed,  $P(B)$  should not be neglected. Given that the incompleteness for open clusters is quite significant at heliocentric distances greater than, say, 1 kpc, we did not currently see this as feasible, cf. Fig. 19.



**Figure 19.** Distance distribution of known open clusters in log-age range  $[7.0, 8.5]$  (blue line with high peak at shorter distances) and of Cepheids at low Galactic latitudes,  $|b| \leq 10^\circ$  (green line). Distances of *bona-fide* CCs from this work indicated by short black lines at the top.

## 5.2 Incompleteness

Despite our aim to maximize the number of clusters and Cepheids considered, only 13 CCs were identified. This is a result of the apparent difference in sample completeness for clusters and Cepheids. Fig. 19 shows the histograms of heliocentric open cluster distances for clusters within the appropriate age range ( $\log_{10}(\text{age}[\text{yr}]) = [7.0, 8.5]$ ), taken from the Dias et al. (2002) catalog as the blue distribution (left peak), and the Cepheid distances compiled (cf. Sec. 3.2.1, green distribution). It is evident that the detection rate of open clusters falls off quickly at distances in excess of 1 kpc, probably since their identification against the field becomes increasingly difficult. Cepheids, on the other hand, are detectable at much greater distances, thanks to their high luminosity and characteristic brightness variations.

However, a cluster’s probability of hosting a Cepheid is governed by stellar evolution and star formation. For a given distribution of stellar masses, only a small fraction will become seen as Cepheids during their lifetime (stars with masses between  $\sim 4-11M_\odot$ ). Due to the nature of the IMF, these intermediate mass stars constitute a small fraction of the total number of cluster stars, and few such intermediate-mass stars are present in typical (small) open clusters. Furthermore, only a fraction (perhaps 10-20 per cent) of a suitable star’s lifetime during the core Helium burning phase is spent on the blue loops, and even less within the instability strip. As a result, few CCs are known. The inverse problem of finding host clusters around known Cepheids, however, can be interesting, e.g. to constrain the survival rate of open clusters.

## 5.3 Implications for the Distance Scale

We are aware of the fact that our calibration, although performed on an ‘optimal’ set, remains inhomogeneous in the true distance moduli employed and the treatment of extinction. It is furthermore based on V-band data obtained in multiple different passbands with varying post processing techniques. These are the main limitations of the data employed. Finally, the number of CCs employed in the fit (10)

is quite small, and the distance to NGC 7790 enters the fit three times, since its three CCs are included in the fit.

We advocate that the ‘optimal’ sample presented, although incomplete (missing host clusters), forms an ideal set for an open cluster-based PLR calibration, since confidence in membership of all clusters was self-consistently evaluated. However, to improve upon the calibration of the PLR, a detailed homogeneous deep photometric study of the host clusters that includes careful treatment of extinction is required. Given the large discrepancies that are found between recent Galactic PLR calibrations (and their zero-points), such an observational campaign would be very desirable.

## 6 CONCLUSION

Focusing on Cepheids, we have performed an all-sky cluster membership census. Our analysis considers an up to 8-dimensional membership-space that includes spatial and kinematic information, as well as parent population parameters (age, iron abundance). Although in some ways limited by data inhomogeneity and incompleteness, we recover most canonical *bona-fide* cluster Cepheids accessible to us and identify three interesting new *bona-fide* CCs.

The newly identified CCs are: SX Car in ASCC 61, S Mus in ASCC 69 (both fundamental-mode pulsators), and V379 Cas in NGC 129 (overtone pulsator). The cluster membership of these three Cepheids must have escaped previous discovery, since the host clusters are not very well-studied, and the Cepheids are located outside the cluster cores, cf. Sec. 4.1.

Based on the available data, we find our results to be inconsistent with membership for the following combinations previously considered in the literature: V442 Car in NGC 3496, TV CMa in NGC 2345, V1726 Cyg in Platais 1, UY Per in Czernik 8, WZ Sgr in Sgr OB7, SZ Tau in NGC 1647, and CS Vel in Ruprecht 79.

Since we can rank candidates through membership probabilities, we consider our *bona-fide* CC sample to be ideal for calibrating the Galactic Cepheid PLR. In Sec. 4.2, we perform this calibration and determine  $M_V = -3.08 \pm 0.50 \log P - 0.94 \pm 0.42$  to be our best solution, which is in chance agreement with the calibration by Sandage et al. (2004). However, this calibration remains limited in precision, due to lack of homogeneity and rather uncertain cluster data. We therefore highlight the need for observational campaigns dedicated to the host clusters of our *bona-fide* CC sample.

The limitations that our work suffers due to inhomogeneous and incomplete data will be significantly reduced in the near future, thanks to the Gaia space mission. Specifically, Gaia will improve our study in the following ways:

- Thousands of new Cepheids (Eyer & Cuypers 2000; Windmark et al. 2011) will be discovered.
- Accurate absolute trigonometric parallaxes of Cepheids will become available up to distances of 6-12 kpc, depending on extinction (Eyer et al. 2012), thereby enabling a direct calibration of the Galactic period-luminosity-relationship similar to the one performed in the seminal paper by Feast & Catchpole (1997). This will remove our partial dependence on existing PLR calibrations when determining cluster membership. Accurate parallaxes to longer period

Cepheids will also be obtained, thereby significantly improving the distance estimates to extragalactic Cepheids.

- The accuracy of proper motions will be improved by orders of magnitude, moving from  $\text{mas yr}^{-1}$  to tens of  $\mu\text{as yr}^{-1}$ , see Lindegren (2010) and the Gaia Science Performance website<sup>16</sup>

- Homogeneous radial velocities (via the RVS instrument) and metallicity estimates (via the spectrophotometric instruments) will be available as membership constraints for a great number of Cepheids.

- Homogeneous metallicity estimates can be obtained through spectro-photometry (Liu et al. 2012).

- Thousands of new open clusters (ESA 2000) will be discovered, forming a more or less complete census of open clusters to distances up to 5 kpc. As was shown in Fig. 19, the distribution of Cepheids is still increasing at these distances. Hence, one may expect to find many more CCs in these parts of the Galaxy.

- Down to magnitude 20, all-sky homogeneous multi-epoch photometry and colors will be obtained that will include all the *bona-fide* clusters mentioned in this work.

- Known clusters will be mapped in unprecedented detail, and intra-cluster dynamics will be accessible to determine membership.

Gaia's data homogeneity will significantly improve error budgets, since no offsets in instrumental zero-points (e.g. in RV) will have to be taken into account. The constraining power in terms of membership will thus be augmented considerably. Correlations between parameters, e.g. proper motion, parallax, and RV, can be determined self-consistently and accounted for (cf. van Leeuwen 2007). Such factors will make the Gaia era particularly exciting for work such as this.

## ACKNOWLEDGMENTS

Many heartfelt thanks are due to the observers who contributed to the Cepheid radial velocity campaign, in particular to: Lovro Palaversa, Mihaly Varadi, and Pierre Dubath. We gratefully acknowledge useful discussions with Maria Süveges, Berry Holl and David G. Turner.

This research has made use of: NASA's Astrophysics Data System Bibliographic Services; the SIMBAD database and Vizier catalogue access tool (cf. A&AS 143, 23), operated at CDS, Strasbourg, France; the International Variable Star Index (VSX) database, operated at AAVSO, Cambridge, Massachusetts, USA; the VO-tool TOPCAT<sup>17</sup>, see Taylor (2005); other online databases that provide Cepheid data, see article body.

## REFERENCES

Akerlof C. et al., 2000, AJ, 119, 1901

An D., Terndrup D. M., Pinsonneault M. H., 2007, ApJ, 671, 1640

Anderson E., Francis C., 2012, Astronomy Letters, 38, 331, (XHIP)

Anderson R. I., Eyer L., Mowlavi N., 2012, in IAU Symposium, Vol. 285, IAU Symposium, pp. 275–277

Bakos G. Á., Lázár J., Papp I., Sári P., Green E. M., 2002, PASP, 114, 974

Balona L. A., Laney C. D., 1995, MNRAS, 277, 250

Baranne A. et al., 1996, A&AS, 119, 373

Baranowski R. et al., 2009, MNRAS, 396, 2194

Barnes, III T. G., Jeffery E. J., Montemayor T. J., Skillen I., 2005, ApJS, 156, 227

Barnes, III T. G., Moffett T. J., Slovak M. H., 1987, ApJS, 65, 307

Barnes, III T. G., Moffett T. J., Slovak M. H., 1988, ApJS, 66, 43

Baumgardt H., Dettbarn C., Wielen R., 2000, A&AS, 146, 251

Benedict G. F. et al., 2007, AJ, 133, 1810

Berdnikov L. N., 2008, Vizier Online Catalog, II/285

Berdnikov L. N., Dambis A. K., Vozyakova O. V., 2000, A&AS, 143, 211

Bersier D., Burki G., Mayor M., Duquenois A., 1994, A&AS, 108, 25

Bono G., Marconi M., Cassisi S., Caputo F., Gieren W. P., Pietrzynski G., 2005, ApJ, 621, 966

Bukowiecki L., Maciejewski G., Konorski P., Strobel A., 2011, Acta Astron., 61, 231, (B11)

Caldwell J. A. R., Keane M. J., Schechter P. L., 1991, AJ, 101, 1763

Camargo D., Bonatto C., Bica E., 2010, A&A, 521, A42

Coulson I. M., Caldwell J. A. R., 1985, South African Astronomical Observatory Circular, 9, 5

Coulson I. M., Caldwell J. A. R., Gieren W. P., 1985, ApJS, 57, 595

Cutri R. M. et al., 2003, 2MASS All Sky Catalog of point sources. NASA/IPAC Infrared Science Archive

Dias W. S., Alessi B. S., Moitinho A., Lépine J. R. D., 2002, A&A, 389, 871, (D02)

Efremov Y. N., 1964, Peremennyye Zvezdy, 15, 242

Eggen O. J., 1980, IBVS, 1853, 1

Eggen O. J., 1983, AJ, 88, 379

ESA, 2000, Gaia: Composition, formation and evolution of the galaxy. Tech. rep., ESA-SCI(2000)4

Evans N. R., 1983, ApJ, 272, 214

Evans N. R., 1992, ApJ, 385, 680

Evans N. R., Welch D. L., 1993, PASP, 105, 836

Eyer L., Cuypers J., 2000, in Astronomical Society of the Pacific Conference Series, Vol. 203, IAU Colloq. 176: The Impact of Large-Scale Surveys on Pulsating Star Research, Szabados L., Kurtz D., eds., pp. 71–72

Eyer L. et al., 2012, Ap&SS, 341, 207

Feast M., 1999, PASP, 111, 775

Feast M. W., 1957, MNRAS, 117, 193

Feast M. W., Catchpole R. M., 1997, MNRAS, 286, L1

Fernie J. D., Evans N. R., Beattie B., Seager S., 1995, IBVS, 4148, 1

Fouqué P. et al., 2007, A&A, 476, 73

Freedman W. L. et al., 2001, ApJ, 553, 47

Froebrich D., Scholz A., Raftery C. L., 2007, MNRAS, 374, 399

Gieren W. P., 1981, ApJS, 46, 287

Gieren W. P., 1985, ApJ, 295, 507

<sup>16</sup> [http://www.rssd.esa.int/index.php?project=GAIA&page=Science\\_Performance#chapter1](http://www.rssd.esa.int/index.php?project=GAIA&page=Science_Performance#chapter1)

<sup>17</sup> <http://www.star.bris.ac.uk/~mbt/topcat/>



- Gieren W. P., Fouqué P., Gomez M. I., 1997, *ApJ*, 488, 74
- Gieren W. P., Matthews J. M., Moffett T. J., Barnes, III T. G., Frueh M. L., Szabados L., 1989, *AJ*, 98, 1672
- Gilmore G. et al., 2012, *The Messenger*, 147, 25
- Glushkova E. V., Koposov S. E., Zolotukhin I. Y., Beletsky Y. V., Vlasov A. D., Leonova S. I., 2010, *Astronomy Letters*, 36, 75
- Gorynya N. A., Irmambetova T. R., Rastorgouev A. S., Samus N. N., 1992, *Soviet Astronomy Letters*, 18, 316
- Groenewegen M. A. T., 2008, *A&A*, 488, 25
- Gupta A. C., Subramaniam A., Sagar R., Griffiths W. K., 2000, *A&AS*, 145, 365
- Harris G. L. H., van den Bergh S., 1976, *ApJ*, 209, 130
- Hoyle F., Shanks T., Tanvir N. R., 2003, *MNRAS*, 345, 269
- Imbert M., 1999, *A&AS*, 140, 79
- Irwin J. B., 1955, *Monthly Notes of the Astronomical Society of South Africa*, 14, 38
- Jaynes E. T., 2003, *Probability Theory - The Logic of Science*. Cambridge University Press
- Kharchenko N. V., 2001, *Kinematika i Fizika Nebesnykh Tel*, 17, 409
- Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., Scholz R.-D., 2005a, *A&A*, 440, 403
- Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., Scholz R.-D., 2005b, *A&A*, 438, 1163, (K05)
- Kharchenko N. V., Piskunov A. E., Schilbach E., Röser S., Scholz R.-D., 2012, *A&A*, 543, A156, (K12)
- Kharchenko N. V., Scholz R.-D., Piskunov A. E., Röser S., Schilbach E., 2007, *AN*, 328, 889
- Kholopov P. N., 1956, *Peremennye Zvezdy*, 11, 325
- King I., 1962, *AJ*, 67, 471
- Kiss L. L., 1998, *MNRAS*, 297, 825
- Klagyivik P., Szabados L., 2009, *A&A*, 504, 959, (KS09)
- Kovtyukh V. V., Soubiran C., Luck R. E., Turner D. G., Belik S. I., Andrievsky S. M., Chekhonadskikh F. A., 2008, *MNRAS*, 389, 1336
- Kukarkin B. V., Kholopov P. N., 1982, *New catalogue of suspected variable stars*
- Lasker B. M. et al., 2008, *AJ*, 136, 735
- Leavitt H. S., Pickering E. C., 1912, *Harvard College Observatory Circular*, 173, 1
- Lindgren L., 2010, in *IAU Symposium*, Vol. 261, *IAU Symposium*, Klioner S. A., Seidelmann P. K., Soffel M. H., eds., pp. 296–305
- Liu C., Bailer-Jones C. A. L., Sordo R., Vallenari A., Borrachero R., Luri X., Sartoretti P., 2012, *MNRAS*, 426, 2463
- Lloyd Evans T., 1980, *South African Astronomical Observatory Circular*, 1, 257
- Lovis C., Mayor M., 2007, *A&A*, 472, 657
- Luck R. E., Andrievsky S. M., 2004, *AJ*, 128, 343
- Luck R. E., Lambert D. L., 2011, *AJ*, 142, 136
- Macri L. M., Stanek K. Z., Bersier D., Greenhill L. J., Reid M. J., 2006, *ApJ*, 652, 1133
- Majaess D. J., Turner D. G., Lane D. J., 2008, *MNRAS*, 390, 1539
- Matthews J. M., Gieren W. P., Mermilliod J.-C., Welch D. L., 1995, *AJ*, 110, 2280
- McSwain M. V., Gies D. R., 2005, *ApJS*, 161, 118
- Mermilliod J. C., Mayor M., Udry S., 2008, *A&A*, 485, 303
- Metzger M. R., Caldwell J. A. R., McCarthy J. K., Schechter P. J., 1991, *ApJS*, 76, 803
- Metzger M. R., Caldwell J. A. R., Schechter P. L., 1992, *AJ*, 103, 529
- Metzger M. R., Caldwell J. A. R., Schechter P. L., 1998, *AJ*, 115, 635
- Perryman M. A. C., ESA, eds., 1997, *ESA Special Publication*, Vol. 1200, *The HIPPARCOS and TYCHO catalogues*. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission
- Petterson O. K. L., Cottrell P. L., Albrow M. D., Fokin A., 2005, *MNRAS*, 362, 1167
- Piatti A. E., Clariá J. J., Ahumada A. V., 2011, *New Astronomy*, 16, 161
- Platais I., 1979, *Astronomicheskij Tsirkulyar*, 1049, 4
- Pojmanski G., 1997, *Acta Astron.*, 47, 467
- Pojmanski G., 2002, *Acta Astron.*, 52, 397
- Pojmanski G., Pilecki B., Szczygiel D., 2005, *Acta Astron.*, 55, 275
- Pont F., Burki G., Mayor M., 1994, *A&AS*, 105, 165
- Pont F., Queloz D., Bratschi P., Mayor M., 1997, *A&A*, 318, 416
- Preston G. W., 1964, *PASP*, 76, 165
- Queloz D. et al., 2001, *The Messenger*, 105, 1
- Queloz D. et al., 2000, *A&A*, 354, 99
- Raskin G. et al., 2011, *A&A*, 526, A69
- Rastorguev A. S., Glushkova E. V., Dambis A. K., Zabolot-skikh M. V., 1999, *Astronomy Letters*, 25, 595
- Robichon N., Arenou F., Mermilliod J.-C., Turon C., 1999, *A&A*, 345, 471
- Roeser S., Demleitner M., Schilbach E., 2010, *AJ*, 139, 2440
- Samus N., Durlevich O., Kazarovets E. V., Kireeva N., Pastukhova E., Zharova A., et al., 2012, *General Catalog of Variable Stars (GCVS database, Version 2012Jan)*, available at the CDS as: *B/gcvs* or under <http://www.sai.msu.su/gcvs/gcvs/>
- Sánchez N., Vicente B., Alfaro E. J., 2010, *A&A*, 510, A78
- Sandage A., 1958, *ApJ*, 128, 150
- Sandage A., Tammann G. A., 2006, *ARA&A*, 44, 93
- Sandage A., Tammann G. A., Reindl B., 2004, *A&A*, 424, 43
- Sandage A., Tammann G. A., Saha A., Reindl B., Macchetto F. D., Panagia N., 2006, *ApJ*, 653, 843
- Ségransan D. et al., 2010, *A&A*, 511, A45
- Soszynski I. et al., 2010, *Acta Astron.*, 60, 17
- Soszynski I. et al., 2008, *Acta Astron.*, 58, 163
- Storm J., Carney B. W., Gieren W. P., Fouqué P., Latham D. W., Fry A. M., 2004, *A&A*, 415, 531
- Storm J. et al., 2011, *A&A*, 534, A94
- Szabados L., 2003, *IBVS*, 5394, 1
- Szabados L. et al., 2012, *MNRAS*, submitted
- Tammann G. A., Sandage A., Reindl B., 2003, *A&A*, 404, 423
- Taylor M. B., 2005, in *Astronomical Society of the Pacific Conference Series*, Vol. 347, *Astronomical Data Analysis Software and Systems XIV*, Shopbell P., Britton M., Ebert R., eds., p. 29
- Tsarevsky G. S., Ureche V., Efremov Y. N., 1966, *Astronomicheskij Tsirkulyar*, 367, 1
- Turner D. G., 1977, *PASP*, 89, 277
- Turner D. G., 1982, *PASP*, 94, 1003
- Turner D. G., 1986, *AJ*, 92, 111
- Turner D. G., 1992, *AJ*, 104, 1865

- Turner D. G., 2010, *Ap&SS*, 326, 219
- Turner D. G., Forbes D., Pedreros M., 1992, *AJ*, 104, 1132
- Turner D. G., Majaess D. J., Lane D. J., Rosvick J. M., Henden A. A. B. D. D., 2010, *Odessa Astronomical Publications*, 23, 119
- Turner D. G., Mandushev G. I., Forbes D., 1994, *AJ*, 107, 1796
- Turner D. G., Mandushev G. I., Welch G. A., 1997, *AJ*, 113, 2104
- Turner D. G., Pedreros M., 1985, *AJ*, 90, 1231
- Turner D. G., Pedreros M. H., Walker A. R., 1998, *AJ*, 115, 1958
- Turner D. G., van den Bergh S., Younger P. F., Danks T. A., Forbes D., 1993, *ApJS*, 85, 119
- Udalski A., Soszynski I., Szymanski M., Kubiak M., Pietrzyński G., Wozniak P., Zebrun K., 1999, *Acta Astron.*, 49, 223
- van den Bergh S., 1957, *ApJ*, 126, 323
- van Leeuwen F., ed., 2007, *Astrophysics and Space Science Library*, Vol. 350, *Hipparcos, the New Reduction of the Raw Data*. Springer
- Vazquez R. A., Feinstein A., 1990, *A&AS*, 86, 209
- Walker A. R., 1985, *MNRAS*, 214, 45
- Walker A. R., 1987, *MNRAS*, 229, 31
- Wilson T. D., Carter M. W., Barnes, III T. G., van Citters, Jr. G. W., Moffett T. J., 1989, *ApJS*, 69, 951
- Windmark F., Lindegren L., Hobbs D., 2011, *A&A*, 530, A76
- Woźniak P. R. et al., 2004, *AJ*, 127, 2436

## APPENDIX A: DETAILS ON INDIVIDUAL CLUSTER CEPHEID COMBINATIONS

### A1 Inconclusive CC Candidates

**A1.0.1 CG Cas - Berkeley 58** CG Cas lies at a separation of  $5.7'$  (outside the core) from Berkeley 58's center, at roughly half the limiting radius. No average cluster RV was found. While the Cepheid's PLR-based parallax is close to that of the cluster and reddening is in agreement, age tends to contradict membership. Furthermore, the differing proper motion ( $\mu_\delta$ ) does not suggest a common point of origin with the cluster and therefore strongly disfavors membership. In light of this, we cannot consider CG Cas a *bona-fide* cluster Cepheid. A detailed investigation of cluster RV and proper motion would be helpful to clarify the membership status.

**A1.0.2 GH Car and Trumpler 18** Membership of the binary overtone Cepheid GH Car (Szabados et al. 2012) in cluster Trumpler 18 was first proposed by Vazquez & Feinstein (1990) and then called into question by Baumgardt et al. (2000) who recommended radial velocity follow-up to draw a firmer conclusion. We compute a high likelihood of 57 per cent based on proper motion, age, and RV. No parallax was calculated, since GH Car is an overtone pulsator. However, the Cepheid's distance listed in the Fernie database (2.2 kpc) is significantly larger than Trumpler 18's (1.4 kpc). Furthermore, the Cepheid is reddened by 0.1 mag more than the cluster, which is consistent with a greater distance to the Cepheid.

Trumpler 18 is sparsely populated (17 members consid-

ered in K05), and there seems to be some uncertainty regarding its age, as seen by the great age difference in K05 and D02 (the latter would be inconsistent with the Cepheid's pulsational age). The average cluster RV is based on the measurements of a single star, and does not agree with that of the Cepheid. Cepheid and cluster are co-moving in proper motion, however, with only a small difference in  $\mu_\delta$ . In summary, there is considerable doubt regarding GH Car's membership in Trumpler 18. A more detailed study of the cluster's parameters, including its average RV would help to clarify this situation.

**A1.0.3 CV Mon and vanden Bergh 1** CV Mon lies right in the center of cluster van den Bergh 1 and was previously considered to be a member of this cluster, cf Turner et al. (1998). The constraints employed in our calculation are parallax, proper motion, radial velocity, and age. From these, we compute a likelihood of merely 4 per cent. The cluster data from B11 is largely inconsistent with the data for the Cepheid, see e.g. parallax (also reddening) and age. The average RV given in D02 is identical to that of CV Mon and no information on the number of stars employed for its determination is given in the original source (Rastorguev et al. 1999); it should thus be discarded as membership constraint. The same goes for the cluster's proper motion which was determined on CV Mon alone by Baumgardt et al. (2000) using the original Hipparcos catalog (hence an offset exists to the proper motion of CV Mon in our data compilation).

Turner et al. (1998) addresses a number of issues encountered when studying this cluster, such as the possible existence of other clusters very nearby van den Bergh 1. An et al. (2007) also noted difficulties involved with this particular cluster, and there may be unreddened field stars present that contaminate the Main Sequence (Preston 1964) impacting distance and age estimates. The cluster's distance estimate by Turner et al. (1998) is in close agreement with the Cepheid's IRSB-distance by Fouqué et al. (2007), and their cluster reddening very close to the Cepheid's  $E(B-V)$  from Kovtyukh et al. (2008). Furthermore their average cluster RV determined on upper main sequence stars also in good agreement with CV Mon's  $v_\gamma$ .

To summarize, two of the four membership constraints compiled are not valid for this particular case, and the determination of distance & age is difficult. The issues encountered here are due to a lack of detail on cluster parameters in the catalogs, especially for a complex case as this, and demonstrate that detailed studies focusing on individual clusters continue to be of great importance.

**A1.0.4 BBSgr and Collinder 394** BBSgr lies at  $21'$  separation from Collinder 394's center, and membership was first studied in detail by Turner & Pedreros (1985) after having been originally suggested by Tsarevsky et al. (1966). Unfortunately, the data available as membership constraints neither really supports, nor excludes, it: The mismatch in RV, although large, is absorbed by the relatively large uncertainties, and parallax, reddening, and age agree. The only real discrepancy is seen in  $\mu_\alpha^*$ , although also this discrepancy barely exceeds the mutual uncertainty.

**A1.0.5 Y Sgr and IC 4725 (M25)** The parallaxes employed (Cepheid  $\pi_{\text{Cep}} = 2.13 \pm 0.29$  from Benedict et al. 2007 and Cluster  $\pi_{\text{Cl}} = 1.61 \pm 0.32$ ) in the calculation nearly agree within their respective error budgets. However, the lower reddening of the Cepheid may make Y Sgr seem to lie in the foreground of M25.

Using the well-established CC U Sgr as a point of reference, we remark that proper motion, age, and metallicity are in excellent agreement between both Cepheids.  $v_\gamma$  differs slightly between the two, which could be explained by the known binarity of U Sgr and Y Sgr. The only discrepancy is in distance, which might be explained by uncertainties in extinction, since M25 lies in the Orion arm (XHIP). Although we calculate a very low prior for this Combo, Y Sgr's distance from cluster center is merely 4.2 pc (projected to M25's distance). From these considerations, it appears that Y Sgr is a could be a member candidate, pending confirmation through follow-up of extinction.

**A1.0.6 BD+47 4225 (GSC 03642-02459) and Aveni-Hunter 1** The star was classified as a Cepheid based on HAT data by Bakos et al. (2002) with a light curve that suggests fundamental-mode pulsation. It lies barely outside  $r_c$  of cluster Aveni-Hunter 1. Unfortunately, proper motion and age are the only available membership constraints available in the literature compiled. Cluster and Cepheid appear to be co-moving in proper motion. Furthermore, the pulsational age of the Cepheid is spot-on with the cluster. However, a very rough distance estimate using the V-band magnitude (10.47) found in the Guide Star Catalog V. 2.3.2 (Lasker et al. 2008) yields a distance of 2.5 kpc for the Cepheid, which is 5 times the cluster distance. Since both objects are thus far not very well-studied, we highlight the need for follow-up of both cluster and Cepheid.

**A1.0.7 Y Car and ASCC 60** Y Car is a double-mode Cepheid in a triple system with a B9.0V companion (Evans 1992) on a known orbit (see the Szabados 2003 binary Cepheids database<sup>18</sup>). The Cepheid lies well inside ASCC 60's projected core. Since the RV of the cluster in K05 was measured on a single star and is identical to Y Car's  $v_\gamma$ , we cannot consider this a valid membership constraint. The age estimates lie just outside each other's error budgets, and reddening agrees. However, the Cepheid's clearly discernible proper motion points to a shorter distance, however, and is in excellent agreement with the proper motion of the cluster. A detailed photometric and radial velocity study of ASCC 60 is required to conclude on the possible membership of Y Car.

**A1.0.8 WZ Car and ASCC 63** Due to the large separation from ASCC 63's core (32') and starkly differing line-of-sight velocities ( $\delta v_{\text{rad}} = 120 \text{ km s}^{-1}$ ), membership would be excluded for this combination, if the probabilities were to be taken at face value.

Looking at the individual constraints, however, we find that age, reddening, parallax (IRSB), and proper motion strongly suggest membership. Suspiciously, the cluster RV is based on only 2 stars, and may therefore not be reliable,

or point towards an ejection event. In any case, this combination remains interesting and follow-up of the cluster is warranted.

**A1.0.9 VZ CMa and Ruprecht 18** This combination yields a likelihood of 83 per cent, since metallicity and age are in excellent agreement between the cluster and the Cepheid. Proper motion is better discernible for the Cepheid than for the cluster, it seems, and reddening for the Cepheid is less strong than for the cluster. The cluster (1.7 kpc) is a fair bit farther than the Cepheid (1.2 kpc, from the Fernie database). VZ CMa may lie in the foreground of Ruprecht 18. However, it would be beneficial to redetermine the Cepheid's distance, e.g. with an IRSB technique, and obtain an average RV for the cluster to better constrain membership for this Combo.

## A2 Combos Inconsistent with Membership

As is clear from the results of our analysis, most cross-matched Combos are inconsistent with membership. We feel it is not necessary to discuss all of these here, and refer the reader to the electronic table. However, a few Combos merit a short description, for instance since they have been considered for membership previously, or lie close to the center of a cluster. These Combos are discussed in the following subsections.

### A2.1 Previously Discussed in Literature

**A2.1.1 SZ Tau and NGC 1647** The membership of SZ Tau in the halo of NGC 1647 was first considered by Efremov (1964) and later studied in more detail by Turner (1992) who concluded that membership is likely. Almost all membership constraints could be applied for this combination. It is conspicuous that parallax and age agree well for cluster and Cepheid (K05), although these are subject to model-dependence and the cluster's age listed in D02 would not agree with SZ Tau's pulsational age. The radial velocities agree within their prescribed uncertainties  $\delta v_{\text{rad}} = 6 \text{ km s}^{-1}$ . Based on proper motion and the very large separation of more than  $2^\circ$ , however, membership can be effectively excluded, as was noted previously by Baumgardt et al. (2000) using (the old reduction of) Hipparcos proper motions.

**A2.1.2 V1726 Cyg - Platais 1** V1726 Cyg's membership results for Platais 1 are based on separation, proper motion radial velocity, and age. The star was first considered for cluster membership by Platais (1979) and Turner et al. (1994). In the available constraints, the only real match between the cluster's and the Cepheid's quantities is in radial velocities. We did not calculate a PLR-based parallax, since V1726 Cyg is an overtone pulsator. However, the Fernie database places V1726 Cyg much closer (2 kpc) than Platais 1 (3.5 kpc, K12), which is consistent with the significantly lower (0.54 mag) reddening of the Cepheid.  $\mu_\delta$  nearly vanishes for the cluster (K12), while V1726 Cyg has clearly detected proper motion in this direction. Finally, the cluster is significantly older than the Cepheid. Taken together, these constraints yield a very low likelihood that indicates non-membership.

<sup>18</sup> <http://www.konkoly.hu/CEP/intro.html>

**A2.1.3 CS Vel and Ruprecht 79** Membership of CS Vel in Ruprecht 79 was thoroughly discussed by Harris & van den Bergh (1976) who credited Tsarevsky et al. (1966) with first suggesting this particular combination. It has since been studied multiple times, e.g. by Walker (1987) and T10. Due to the sparse nature of the cluster, its reality as such was doubted by McSwain & Gies (2005), who conclude that Ruprecht 79 rather be a hole in the dust of the Sagittarius-Carina spiral arm than a physical open cluster.

The data compiled in this work would place the Cepheid about 1 kpc farther from the Sun than Ruprecht 79 (with E(B-V) by 0.1 mag higher for the Cepheid than the cluster), while the average cluster RV (K05) is not determined clearly enough to provide a strong constraint. We furthermore note that proper motion for the cluster and the Cepheid strongly contradict membership. Therefore, even if Ruprecht 79 is, indeed, an open cluster, the information available would be inconsistent with membership, yielding a likelihood of merely 5 per cent.

**A2.1.4 V442 Car** has previously been considered for membership in NGC 3496 by Balona & Laney (1995) who concluded it to be a background star, based on age and reddening. From proper motion and separation, we come to the same conclusion. Furthermore, a rough distance estimate for a 14<sup>th</sup> magnitude 5.5 d Cepheid excludes membership in a cluster located approx. 1 kpc from the Sun.

**A2.1.5 UY Per and Czernik 8 or King 4** Turner (1977) suggested that UY Per could be a member of Czernik 8 or King 4. Turner et al. (2010) again mentions the latter combination. Our results, however, are inconsistent with membership in either cluster, based on the constraints parallax, proper motion, and age.

The ‘likelier’ of the two Combos is Czernik 8, for which parallax and age are in relatively good agreement; the Cepheid is slightly farther away and has larger reddening. However, kinematically, the cluster’s vanishing proper motion (K12) would be inconsistent with membership of the rather fast moving Cepheid ( $\mu_\alpha^* = -6.15 \pm 2.8 \text{ mas yr}^{-1}$ ,  $\mu_\delta = 12.89 \pm 2.9 \text{ mas yr}^{-1}$ , PPMXL).

**A2.1.6 TVCMa and NGC 2345** The membership constraints parallax, proper motion (vanishes), and age, agree within their respective uncertainties. Radial velocity differs by  $\sim 20 \text{ kms}^{-1}$  between cluster and Cepheid. In line with the prior, we take this to indicate non-membership for this Combo.

## A2.2 Foreground Cepheids

**A2.2.1 OPCas and AUCas in CasOB7** Although the parallaxes compiled for OPCas and CasOB7 agree within their uncertainties, the large difference ( $\sim 1 \text{ mag}$ ) in reddening is indicative of non-membership. No Cepheid RV is available, and proper motion vanishes. Although the age agrees well, we cannot have confidence that the cluster parameters are correct, e.g. if reddening was overestimated. The same is true for AUCas.

**A2.2.2 CO Aur and Kaposov 12** All available information points towards a foreground Cepheid: the literature distance to CO Aur is significantly shorter than to Kaposov 12, and proper motion is non-vanishing for the Cepheid, while it vanishes for the cluster.

## A2.3 Background Cepheids

There are a few Combos worth mentioning here that yield high likelihoods ( $> 75$  per cent) and low priors ( $\sim 0$  per cent). These ones are WX Lac and IC 1442, GM Cas and IC 1805, as well as X Vul and Turner 9. The cluster and Cepheid parallaxes for each of these Combos agree within the stated uncertainties. However, for all of these Combos, the Cepheid reddening is significantly (at least twice, or 0.65 mag) larger than for the cluster. Larger reddening is indicative of greater distance. We therefore dismiss these candidates, taking into account also the remaining membership constraints.

**A2.3.1 V410 Vul and NGC 6823** Located outside  $r_{\text{lim}}$ , proper motion vanishes for the Cepheid, but not for the cluster. The faintness of Cepheid (18<sup>th</sup> mag) hints at a location in the background of NGC 6823.

**A2.3.2 NSV 19957 and NGC 5281** The small amplitude Cepheid (given 2<sup>nd</sup> rank as a Cepheid candidate by Caldwell et al. 1991) lies outside  $r_c$  in the vicinity of NGC 5281. Although proper motion yields a high likelihood (0.89) and the separation is rather small in absolute terms (6.9’), the Cepheid is likely a background object. The mean I-band magnitude of 13.18 appears too faint for NGC 5281’s distance of just above 1 kpc. We conclude that membership is unlikely, but suggest follow-up of the Cepheid to better understand its nature.

**A2.3.3 NSV 19262 and Ruprecht 101** The star was deemed a good Cepheid candidate by Caldwell et al. (1991) and lies at a separation of  $4.4' > r_c$  from Ruprecht 101’s center. The faint I-band magnitude of approx. 13.15 together with a period of roughly 15 d (estimated from Fig. 7ff in the original report) expose the star’s nature as a background object to the 2.3 kpc-distant cluster.

**A2.3.4 NSVS 2089061 and NGC 1496** Cepheid and cluster appear to be co-moving and have matching ages, with the Cepheid located in the cluster’s limiting radius. More information on the Cepheid is required to properly assess membership. However, from a rough distance estimate it would seem likely that the Cepheid is located at around 3 to 5 kpc, whereas the cluster is much closer (1.4 kpc).